

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 21, 2002, 14:49:40 ; Search time 42.14 Seconds
(without alignments)
3642.714 Million cell updates/sec

Title: US-09-729-653-2
Perfect score: 7257
Sequence: 1 HSLIGRCSRGILGDGNVAVC.....AAOLKISDCTQYLRTHEEL 1382

Scoring table: BLASTSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues
Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: A.Geneseq_032802.*
2: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1980.DAT.*
3: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1981.DAT.*
4: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1982.DAT.*
5: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1983.DAT.*
6: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1984.DAT.*
7: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1985.DAT.*
8: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1986.DAT.*
9: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1987.DAT.*
10: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1988.DAT.*
11: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1989.DAT.*
12: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1990.DAT.*
13: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1991.DAT.*
14: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1992.DAT.*
15: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1993.DAT.*
16: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1994.DAT.*
17: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1995.DAT.*
18: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1996.DAT.*
19: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1997.DAT.*
20: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1998.DAT.*
21: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1999.DAT.*
22: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2000.DAT.*
23: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2780	38.3	1366	22	ABB65570	Drosophila melanog
2	2623	36.1	515	22	AAAB80287	Human protein SEO
3	2544	35.1	502	22	AAAF93503	Human protein SEO
4	1229	16.9	244	22	AAU14657	Novel bone marrow
5	1163	16.0	237	22	AAAG74711	Human colon cancer
6	706	9.7	138	22	AAU14751	Novel bone marrow
7	298	4.1	109	22	AAU33123	Novel human secret
8	293	4.0	1749	22	AAAT91330	Human protein SEO
9	289	4.0	1205	22	AAAB58628	Drosophila melanog
10	274.5	3.8	72	22	AAU33013	Novel human secret
11	263	3.6	1241	22	AAAT9062	Human protein SEO

12	260.5	3.6	1252	21	AAAB43197	Human ORFX ORF2961
13	259.5	3.6	1690	21	AAAB42248	Human ORFX ORF2012
14	235.5	3.2	235.5	22	AAE01339	Human gene 24 enco
15	235.5	3.2	993	22	AAE93308	Human protein sequ
16	235	3.2	1443	22	AAE68472	Drosophila melanog
17	216.5	3.0	864	22	AAAB94031	Human protein sequ
18	213	2.9	942	22	AAAB80046	Human protein sequ
19	206.5	2.8	873	22	AAAB95245	Human protein sequ
20	206	2.8	821	22	AAAB93491	Human protein sequ
21	174.5	2.4	968	22	AAAB63037	Drosophila melanog
22	170	2.3	887	22	AAE68631	Drosophila melanog
23	167	2.3	1332	20	AAAY01084	Human IKAP protein
24	167	2.3	1332	22	AAAB66345	Human IKAP. Homo
25	166	2.3	1332	22	AAAB93160	Human ORFX ORF3013
26	164	2.3	496	21	AAAB43249	Human polyprotein
27	161.5	2.2	1246	22	AAAB39120	Human polyprotein
28	157	2.2	845	21	AAAC43631	Arabidopsis thalia
29	157	2.2	923	21	AAAC43630	Arabidopsis thalia
30	156.5	2.2	2998	18	AAAI9691	ATM mutant 3403del
31	155	2.1	1270	22	AAAB99141	Pyrococcus abyssi
32	155	2.1	1455	22	AAAB99140	Pyrococcus abyssi
33	155	2.1	1458	22	AAAB96825	Putative P. abyssi
34	154.5	2.1	2386	18	AAAI3153	S. pombe Rad3 poly
35	154	2.1	769	22	AAAB13410	Novel human diago
36	150.5	2.1	1294	22	AAAB40906	Human polyprotein
37	149.5	2.1	475	21	AAAY53985	Amino acid sequenc
38	147.5	2.0	1149	22	AAAY97640	Apat-1XL/Delta2-10
39	147.5	2.0	1237	22	AAAY97638	Apat-1MD protein s
40	147.5	2.0	1248	22	AAAY97636	Apat-1XL protein s
41	147.5	2.0	1248	22	AAAY97639	Apat-1XL-M368L pro
42	147.5	2.0	1248	22	AAAY97641	Apat-1XL-L10A prot
43	147.5	2.0	1248	22	AAAY97642	Apat-1XL-D27A prot
44	147.5	2.0	1248	22	AAAY97643	Apat-1XL-E390 prot
45	147.5	2.0	1248	22	AAAY97644	Apat-1XL-R630 prot

ALIGNMENTS

RESULT 1	
ABB65570	standard; Protein; 1366 AA.
XX	
XX	ABB65570:
AC	
XX	
DT	26-MAR-2002 (first entry)
XX	
XX	
DE	Drosophila melanogaster polypeptide SEQ ID NO 23502.
XX	
XX	
KW	Drosophila: developmental biology; cell signalling; insecticide;
KM	pharmaceutical.
XX	
OS	Drosophila melanogaster.
XX	
PN	WO200171042-A2.
XX	
PD	27-SEP-2001.
XX	
PF	23-MAR-2001; 2001MO-US09231.
XX	
PR	23-MAR-2000; 2000US-191637P.
XX	
PR	11-JUL-2000; 2000US-0614150.
XX	
PA	(PEKE) PE CORP NY.
XX	
PI	Venter JC, Adams M, Li PWD, Myers EW;
XX	
DR	WPI: 2001-656860/75.
XX	
DR	N-PDB; ABL09673.
XX	
PT	New isolated nucleic acid detection reagent for detecting 1000 or more
PT	genes from Drosophila and for elucidating cell signalling and cell-cell
PT	interactions -

DAVIS M.
09/1724653
Seq. ID 2

```

XX PS Disclosure: SFO ID NO 23502; 21bp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signaling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
CC sequences (AB101840-AB16175) and the encoded proteins.
CC (AB57737-AB572072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1366 AA;

Query Match      38.3%; Score 2780; DB 22; Length 1366;
Best Local Similarity 40.8%; Pred. No. 7, 6e-241;
Matches 553; Conservative 286; Mismatches 468; Indels 50; Gaps 14;

OY 57 IGFAMQKTSIGNVLAATGADYIVKIFEDRHGQKSEINLPNCVAMDMKDGDLAVIAEKS 116
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 19 vyfiwqka--llattgdsvalynrgqlvqrlilsglscgfawdgdglilgtssg 75

OY 117 SCITVMDAFTNKTSOLDNGMRDOMSFLMSKVSFLAVGTGKLNKTYNHQT-SRKIPVL 175
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 76 pulctwdysngkisevtelrldplctilwskqqlavtgrfgnailynhssgkrlpt 135

OY 176 GKHTRTTCGCNNAENLXALGEGDKMITVSNOEGDTRTOYVRSPPXNQOFLMKRDDPT 235
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 136 ghsrkrttcgawsagqllalgedsksfisnedgdlvrvqrlrdaptamyfaemanderi 195

OY 236 SAAESMISVVLKKTFLFNLNEPDNPADLEFQDGNITVCYNYWYGDGRIMTGFSCGHEF 295
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 196 -agdaismilgkrtfllylpepenetelgfsrsgyslmqhkfwfdgyillgfsnghv 254

OY 296 VISTHTGELGGEIFOARNKDNLTISAVSQTILNKVATCGDNCKIKIDVLDMVYIWL 355
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 255 aistpckdvqgelwqkmdsklclgacpldiavscgddsklhtlnlqetercliv 314

OY 356 DEENKGLTISWTDOQLALSTORCSLHVFLTKLPILDACSRTIAYLSLEVT--- 412
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 315 -pdhagvqmldspdgqllavetlnhtylylvtklphlavesaprlivllsilaeslyiy 373

OY 413 ANPVEGELPITYSDVEPPEFVAVGLYHLAVGNMNNRAFPVLCEN--AVKKIKMEYLCGT 469
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 374 apdtktslprfprlegelrtimavpynfatglekhwfydglkslgeepriiserdfrs 433

OY 470 VASICLHSDYAALFEGKVQLHISELIDAOEERT---RLFPVDDCKRLCHALITSD 526
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 434 vesmlinaaycaalcipqlllqalaadnpckklqavfrtalpmmpsdavltcltslq 493

OY 527 FLIVGTDTGVVOFYIEDMVFVNDYRHPYVKRIFFPDNGTRLVTFIDEKSDGFVYCPVND 586
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 494 llfatdighlvtfelekdsctlyrhsmgtrqlfmdigtckvifidhsgsyvflprve 553

OY 587 ATVEIPDSPTTKGVLMENKMPDKGVFLAYDDDKYTTYFHKDTIQGAKVILAGSTKYPE 646
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 554 eaallipdpikgclglwd--lcpnlflisydarivnthfvfshvsgtchtlmvgesklmp 611

OY 647 AHKPLLLNGELTCQTOSGKVNNTYLSHGFLSLNKDKSPDLRLPMALNMLKRRSDMW 706
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 612 gqfplllcgemaalhldggqyatqstlth--vvpsnsqaanlqnl---lllrydeay 665

OY 707 EMCRLINDAAMNELARACLHMEVEFAIRYRRIGNVIVSLPOIKIGIEDYVNLACHL 766
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 666 klckmqnsgsawrefgqaisledpldairyqlgdaalvalnaselyvelldmlngc 725

OY 767 AMFTNDVLAODLYLASSCPALAEKMRDLQHWDSALQAKLAPDQITPFSIKVAYIDG 826
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 726 ctllagydqakehmlkytyraatldrcdlilqwdqalllahnkdpgevpylareyaqgle 785

```

```

OY 827 FAGDYVNALEHKEGITTG-----NKEHDEACTAGAAQMSIRMGDIR 868
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 786 fngnydalhyekykedlinsketetaldmspspeehvrlckmgjartsaigrdfr 845

OY 869 RGVNQLKTHPSRVLKNDGCALLENMKQFSPAQLYKSGLYDKAASVYTRSKMMAVVGSL 928
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 846 rlgqvaveledqllldccellatvghlteaaglyer99rfydeagcghyalmmvannai 905

OY 929 LPHVSSPKIHLQYAKAREADGRYKREAVVAYENAKQMSVIRYLDHLMNPERAVNIVRET 988
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 906 lpkvsktllhaayakakendgnyeeaisryrlagldidacvrylidlcdphaaselvles 965

OY 989 QSLDGAKMVARFRLQDYGSAIQFLVMSKCNNEATLQAQHKMKMTYADITGS-EDTN 1047
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 966 rsmdsakllaktyklygdvqelqflivgcveeaalagrhmkrlrhgelleryenaks 1025

OY 1048 EDVOSIALYEEGKRYLQKGFLLCGOYSRALKHFLKCP--SSEDNVAVIEMAIETVGOA 1105
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1026 sdylalahyfegekyllyagkyflaretktalrlllkasafmneeqvslstalcvats 1085

OY 1106 KDELTLNOLIDLHLENDGMPKDAKYLRLYALOKYREAAQTALIIAREEOSAGNYRNA 1165
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1086. nneqatqliefligevdvpkprylfrlymarkhykdaaktavliandeqiaagnykaa 1145

OY 1166 HDVLPSTVAELSSQKIKIPSEVATNMLHSTYLKVIHKNQGDHMKGAKMLRVANNISK 1225
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1146 rdllysmvaygelrrmlsvtaemrhqfillhyrtylvirhvkllgnhllaakllqvaaclsq 1205

OY 1226 PFSHVPIPLTIVYIECHRGLKNSAFSFAAMLMPREYRSKIDAKYKKLEGVWRPDP--- 1282
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1206 fhelipiltsviechragllksaflyastlmrpdryqldipryakklieslyrkakpi 1265

OY 1283 ---ISEIEEATTPCEPKFLPECELLCPGCKNSIPYCIATGRHMLKMDWTVCPHCDFPA 1339
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1266 kglrdeidgmetcpcpdcnsalnmewtvcyscktlpicataghlhikqlmtscpqdflc 1325

OY 1340 LYSELKIMLNTESGTCMCSERLMAOLKKISDCTQL 1376
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1326 fraemenllisengcpcmgenvapeqllidvedirpyl 1362

RESULT 2
AAM80287
ID AAM80287 standard; Protein; 515 AA.
XX
AC AAM80287;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human protein SEQ ID NO 3933.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation.
OS Homo sapiens.
XX
PN WO200157190-A2.
XX
PD 09-AUG-2001.
XX
PF 05-FEB-2001; 2001WO-US04098.
XX
PR 03-FEB-2000; 2000US-0496914.
PR 27-APR-2000; 2000US-0560875.
PR 20-JUN-2000; 2000US-0598075.
PR 19-JUL-2000; 2000US-0620325.
PR 01-SEP-2000; 2000US-0654936.
PR 15-SEP-2000; 2000US-0663561.
PR 20-OCT-2000; 2000US-0693325.
PR 30-NOV-2000; 2000US-0728422.

```

(HYSE-) HYSEQ INC.
 Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
 Zhao Qa, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
 Xue AJ, Yang Y, Wejhrman T, Goodrich R;
 MPI: 2001-476283/51.
 N-PSDB: AAK53420.
 DR
 XX
 Nucleic acids encoding polypeptides with cytokine-like activities,
 useful in diagnosis and gene therapy -
 PT
 PS
 Claim 20: Page 471; 6221pp; English.
 XX
 The invention relates to polynucleotides (AAK51456-AAK53435) and the
 CC encoded polypeptides (AAM78323-AAH80302) that exhibit activity elating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polypeptides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, hematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation.
 CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
 CC (AAM80020) are omitted as the relevant pages from the sequence listing
 CC were missing at the time of publication.
 CC
 XX
 Sequence 515 AA:
 SQ

Query Match	Similarity	Score	DB	Length
Best Local	99.6%	Pred. No. 1.9e+27;		
Matches	508:	Conservative	1;	Mismatches 1; InDels 0; Gaps 0
QY	797	QHMDSALQIAKLAAIDQDIPFISKEVAYIQLEFAGDYYVNAALAHYKGTGGDNKEHDEACLAG	856	
DB	6	eimwsaqlakhlapdqqlpfskexyafqfetsgdyynalabhekyltglnkhenecslag	65	
QY	857	VAQMSIRMGDILRRGVNALKHPRSVLKRDGCAILENMKOFSEPAOLYEGGLYYDKAAY	916	
DB	66	vagmslmgdlirryngqalnhpsrlykrcdgaallnmkqfseaaqlyekgllydkaasy	125	
QY	917	IRSKMAVYGDILPHYSSPKIHLOYAKKKEADGRKKEAVVAAVENMKOMOSVIRIYDHLN	976	
DB	126	lrskmwakgdlilphvsespklhlygakakeaagrykxaaaenaqwsvirlyrlydhlh	185	
QY	977	NPEKAVNIYREIQSLDGAHVAFELQLEDYCSALQELFVMSKCNMNAFLLAODHNMELY	1036	
DB	186	npekeavnliretqslidgakmmarfflqldgdyssalqflvmskcnneaflaagqhmkmely	245	
QY	1037	ADITIGSEDTNEDYQSIALYFEGEKRYLOAGKFFLLCGOYSRALKHFLKCPSEEDVAIE	1096	
DB	246	aditigsedctnedysqisialyfegekrylbaqgffllcgqsyralxhflkcpssedvnaie	305	
QY	1097	MAIETPYGAKDELTLTQQLIDHLLGENDGMKPAKTYFLRYMLKORYREAKOAPLITAREE	1156	
DB	306	maietpygqakdeltlrqqlldhlhgendgmprkakyflrymlakgyreaaqtaillaree	365	
QY	1157	QSGANVRNAHDVLFMSYAEILKSQIKIPSEMAITNLILHSYLLVYKIHVNGDHMKGARM	1216	
DB	366	qsgangymahdvlfmsyaeelksqklkpsemaatnlmlhsyllvkiahvngdhmkgarml	425	
QY	1217	IRVANNISKPSSHIVPILTSTVIECHRAGLKNSAFSFAAMLMRPYRSKIDAKYKKKIBG	1276	
DB	426	irvanmiskfpsihlypilstvtechrqglknsafsfaamlmrpyrskidakykkkieg	485	
QY	1277	MVRRPDISIELEATTPCPFCFKFLPRCELL	1306	
DB	486	mvrtpdisieleattpcpcfkflprsell	515	

RESULT 3
AAM79303
ID AAM79303 standard; Protein; 502 AA.
XX
AC AAM79303;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human protein SEQ ID NO 1965.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation.
XX
OS Homo sapiens.
XX
PN MO200157190-A2.
XX
PD 09-AUG-2001.
XX
PE 05-FEB-2001; 2001MO-US04098.
XX
PR 03-FEB-2000; 2000US-0496914.
XX
PR 27-APR-2000; 2000US-0560875.
XX
PR 20-JUN-2000; 2000US-0598075.
XX
PR 19-JUL-2000; 2000US-0620325.
XX
PR 01-SEP-2000; 2000US-0654936.
XX
PR 15-SEP-2000; 2000US-0663561.
XX
PR 20-OCT-2000; 2000US-0693325.
XX
PR 30-NOV-2000; 2000US-0728422.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Dymnac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AD, Yang Y, Wejrtman T, Goodrich R;
XX
DR WPI: 2001-476283/51.
DR N-PSDB; AAK52436.
XX
PT Nucleic acids encoding polypeptides with cytokine-like activities,
PT useful in diagnosis and gene therapy -
XX
PS Claim 20; Page 4376-4377; 6221pp; English.
XX
XX The invention relates to polynucleotides (AAK51456-AAK53435) and the
XX encoded polypeptides (AAM78373-AAK80302) that exhibit activity elating t
XX cytokine, cell proliferation or cell differentiation on which may induce
XX production of other cytokines in other cell populations. The
XX polynucleotides and polypeptides are useful in gene therapy, vaccines or
XX peptide therapy. The polypeptides have various cytokine-like activities,
XX e.g. stem cell growth factor activity, haematopoiesis regulating
XX activity, tissue growth factor activity, immunomodulatory activity and
XX activin/inhibin activity and may be useful in the diagnosis and/or
XX treatment of cancer, leukaemia, nervous system disorders, arthritis and
XX inflammation.
XX Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
XX (AAM800020) are omitted as the relevant pages from the sequence listing
XX were missing at the time of publication.
XX
XX Sequence 502 AA;

	Query Match	Similarity	Score	DB	Length
Best Local	35.1%	99.0%	Pred. No. 2.5e+20		502
Matches	493	Conservative	2	Mismatches	3
				Indels	0
				Gaps	0

QY	708	MGRITLDEANWELARACIHHMEVEFAIRYVRIRRGVGVMSLEQIKGIEDYNNLAGHIA	767
DB	1	mcrlindeaanwelaacchhmevefaivrrirvgvrvmsleqikgiedynllagha	60

```

QY 768 MFTNDYNIQAODLYLASSCITAALEENRDLQHMDSALQAKHLAPDOIPIITSEYAIQLEF 827
  |||||||
DB 61 mftndynlaqdiyljasscpiaaleemrtdlqhwdsalqlakhlapdqjpfitskeyaqllef 120
QY 828 ACDDYNALAHYKSGITGDNKHEDEACLAGVAOMSTIRMGDIRRGVNOALKHPSVYLKRDGC 887
  |||||||
DB 121 agdynalalahykgitgdnkheadeaclagvaqmslimgdlrrgvnqalkhpsrvlkrdcg 180
QY 888 ALLENKMGFSEAAQLYEKLKYDKASVYIRSKNNMAKVDLLPHVSSPKIHQYAKAKKA 947
  |||||||
DB 181 allemkmgfseaaqlyeklgydkasvylrsknnakvgdllphvsspkihqyakakka 240
QY 948 DGRYKAAVAYENAKWQSVIRIYIDLHNNPEKAVIVETQSLGCAKWAARFLOLQGY 1007
  |||||||
DB 241 dgrkvaavayenakwqsviriyidhlnpekavivretqslgdakmwarfflqlqgy 300
QY 1008 GSAIFPLVSKCNNEAFTLAQOQHNKMEIYADIGSEDTNEDQSFALFEEGKRYLQAG 1067
  |||||||
DB 301 gsaifplvmskcnneafllaqghnkmeiyadlgsedtnedqsfalyfegekrylqag 360
QY 1068 KFFLLCGQYSRALKHFLKCPSEEDNVAIEMALETVGQARDELTLNQLIDLHLEGENDMPK 1127
  |||||||
DB 361 kffllcgysralkhflkcpseedenvalemaletvgqakdelldlqdlhlllgendsmkp 420
QY 1128 DAKYLFRLYMAKQYREAQTAIITAREEOSAGNRYNADVLFSWYAELEKSKIRIPSEM 1187
  |||||||
DB 421 dakylfrylmalqyreaqtailltareeqsagynrnahdvlfswaelksqkikripsm 480
QY 1188 ATNLMITHSYITLVKIHVK 1205
  |||||||
DB 481 atnlmlthsyilvtrfmk 498

```

RESULT 4
 ID AU14657
 AU14657 standard; Protein: 244 AA.

```

AC AU14657;
DT 24-OCT-2001 (first entry)
DE Novel bone marrow polypeptide #56.
KM Bone marrow; diagnostic; therapeutic; gene therapy; antigenic;
  haematopoiesis; myeloid; lymph cell disorder; tissue regeneration;
  wound healing; nutritional supplement; immune disorder;
  severe combined immunodeficiency; SCID.
OS Homo sapiens.
PN WO200157187-A2.
PD 09-AUG-2001.
PF 05-FEB-2001; 2001WO-US03782.
PR 03-FEB-2000; 2000US-0496914.
PR 20-JUN-2000; 2000US-0598075.
PR 19-JUL-2000; 2000US-0620325.
PR 30-NOV-2000; 2000US-0250683.
RA (HYSE-) HYSEQ INC.
PI Ford JE, Boyle BJ, Tang YT, Liu C, Asundi V, Zhou P, Xue AJ;
  Ren F, Drmanac RT;
XX WPI: 2001-488875/53.
XX N-PSDB: AAS22862.
XX Nucleic acids encoding bone marrow polypeptides, useful in diagnostic
  and gene therapy -
PS Claim 10; Page 239-240; 392pp; English.

```

```

XX AAU14602-AAU14794 represent novel bone marrow polypeptides of the
CC invention. The proteins and corresponding coding sequences may be used
CC in the prevention, diagnosis and treatment of diseases associated with
CC inappropriate bone marrow polypeptide expression. For example, to treat
CC disorders associated with decreased expression by rectifying mutations
CC or deletions in a patient's genome that affect the activity of the
CC polypeptides by expressing inactive proteins or to supplement the
CC patient's own production of the polypeptides. Additionally, the nucleic
CC acids may be used to produce the polypeptides, by inserting the nucleic
CC acids into a host cell and culturing the cell to express the protein.
CC The nucleic acid and its complementary sequences may also be used as DNA
CC probes in diagnostic assays to detect and quantitate the presence of
CC similar nucleic acid sequences in samples, and therefore which patients
CC may be in need of restorative therapy. The proteins may also be used as
CC antigens in the production of antibodies against bone marrow proteins
CC and in assays to identify modulators of their expression and activity.
CC The anti-bone marrow protein antibodies and antagonists may also be used
CC to down regulate expression and activity. The antibodies may also be used
CC as diagnostic agents for detecting the presence of the protein in samples
CC (e.g. by enzyme linked immunosorbent assay (ELISA)). The proteins
CC may be used to regulate haematopoiesis activity, and consequently in the
CC treatment of myeloid or lymph cell disorders; in tissue regeneration,
CC such as wound healing; as a nutritional supplement; and in treatment of
CC immune disorders such as severe combined immunodeficiency (SCID).
SQ Sequence 244 AA:

```

Query Match 16.9%; Score 1229; DR 22; Length 244;

Best Local Similarity 96.3%; Pred. No. 4e-102; Matches 234; Conservative 0; Mismatches 3; Indels 6; Gaps 1;

```

QY 100 MDMDKGDVLAIVAEKSSCIYLDANTNTSOLDNMGMDQMSFLMSKVGSFLAVGTVGK 159
  |||||||
DB 1 mdmdkdgdvlaiviaekssciylwdantntksqldnmgdmqmsfllmskvgsflavgtvgk 60
QY 160 NLXIYNHOTSRKIPVLGKTKRITTCGNNAEULXALGEGDKMITVSNQEGDITRTQVNS 219
  |||||||
DB 61 nllyhngtsrkilpylgktkrltcgcwaaenllaigedkmltvsnegedtlrtqgvs 120
QY 220 EFXNNOFLMKDDPRTSAESMISVVLGKTLFPLNLNPPDPADPEFOQDGNF----- 274
  |||||||
DB 121 efxnnofllmkndrttsaesmsvylgkllflnlhnpdpadefqdgfnlvcynw 180
QY 275 -VCYNNWYGGRIMIGFSCGFVYVISTHTGELGOELFQARNHKNLSTLAVSQTLNKVATC 333
  |||||||
DB 181 yvcynwyggrlmigfscghfvvsthtgelnqelqfarnhknldtstlavsgtlntkvatc 240
QY 334 GDN 336
  |||
DB 241 gdn 243

```

RESULT 5
 ID AAG74711
 AAG74711 standard; Protein: 237 AA.

```

AC AAG74711;
DT 03-SEP-2001 (first entry)
DE Human colon cancer antigen protein SEQ ID NO:5475.
XX Human; colon cancer; colon cancer antigen; diagnosis; detection;
  colorectal carcinoma; chromosome 1.
XX Homo sapiens.
XX WO200122920-A2.
XX 05-APR-2001.

```



```

PF      28-SEP-2000; 2000MO-US26524.
XX
XX      29-SEP-1999;   99US-0157137.
PR      03-NOV-1999;   99US-0165280.
XX
XX      (HUMA-) HUMAN GENOME SCI INC.
PA
XX      Ruben SM, Barash SC, Birse CE, Rosen CA;
PI      WPI: 2001-235357/24.
XX      N-PSDB; AAH34116.
DR
XX      Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
PT      useful for preventing, diagnosing and/or treating colorectal cancers -
XX
XX      Claim 11; Page 7076-7077; 9803pp; English.
PS
XX      AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
CC      cancer-associated nucleic acid molecules (N) and proteins (P), where
CC      the proteins are collectively known as colon cancer antigens. The colon
CC      cancer antigens have cytoskeletal activity and can be used in gene
CC      therapy and vaccine production. N and P may be used in the prevention,
CC      diagnosis and treatment of diseases associated with inappropriate P
CC      expression. For example, N and P may be used to treat disorders
CC      associated with decreased expression by rectifying mutations or deletions
CC      in a patient's genome that affect the activity of P by expressing
CC      inactive proteins or to supplement the patients own production of P.
CC      Additionally, N may be used to produce the colon cancer-associated Ps,
CC      by inserting the nucleic acids into a host cell and culturing the cell
CC      to express the proteins. N and P can be used in the prevention, diagnosis
CC      and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
CC      and AAB77789 represent sequences used in the exemplification of the
CC      present invention.
CC      N.B. Pages 666 to 682 and page 7053 of the sequence listing were
CC      missing at time of publication, meaning no sequences are present for
CC      SEQ ID NO:1027 to 1052, 7921 and 7922.
XX
S0      Sequence    237 AA;

Query Match          16.0%; Score 1163; DB 22; Length 237;
Best Local Similarity 99.1%; Pred. No. 3.4e-96;
Matches 228; Conservative 1; Mismatches 1; Indels 0; Gaps 0

QY      1050 YQSIALYFEGEKERYLQAQKFFLLCGOYSRALKHFLKCPSSSEDNVALEMAIEFTGAKDEL 1109
DB      1 ygsialylegeekerylqagffllcgysralkhflkcpssedvalemaletevgaqkdell 60
QY      1110 LTNOILDIHLGENDGMKPKDAKTLFRLYMALKQYREAAQTALITIRREGDSAGRNRNHADV 1169
DB      61 ltnqildihlgendgmprkakylfclymalkyreaaqtalilaareqragynrhadv 120
QY      1170 FSMVRELKSQKIKRISSEMATNLMLHSYLYLVRIHVNGDGHMGARMLIIVANNISKFPSPH 1229
DB      121 fsmvelksqklkipsematnlmlhsyllvkilhvngdghmgarmllivanniskfpsh 180
QY      1230 IVPILTSTVIECHRRAGLKNSAFSPAAMLMRPYRSKTIDAKKKYKKEIGMVR 1279
DB      181 ivpiltstviechrraglknsafspaamlmrepyrsktidakkykkklegmwg 230

RESULT        6
ID      AAU14751 standard; Protein; 138 AA.
XX      AAU14751;
AC
XX      AAU14751;
DT      24-OCT-2001 (first entry)
XX
DE      Novel bone marrow polypeptide #150.
XX
KW      Bone marrow; diagnostic; therapeutic; gene therapy; antigenic;
KM      haematopoietic; myeloid; lymph cell disorder; tissue regeneration;
```

KW	wound healing; nutritional supplement; immune disorder;
KM	severe combined immunodeficiency; SCID.
XX	
OS	Homo sapiens.
PN	MO200157187-A2.
XX	
PD	09-AUG-2001.
XX	
PF	05-FEB-2001; 2001WO-US03782.
XX	
PR	03-FEB-2000; 2000US-0496914.
XX	
PR	20-JUN-2000; 2000US-0598075.
XX	
PR	19-JUL-2000; 2000US-0620325.
FR	30-NOV-2000; 2000US-0250683.
XX	
PA	(HYSE-) HYSEQ INC.
PI	Ford JF, Boyle BJ, Tang YT, Liu C, Asundi V, Zhou P, Xue AJ;
PI	Ren F, Dermanac RT;
XX	
DR	WPI: 2001-488875/53.
XX	
DR	N-PsDB: AAS23056.
XX	
PT	Nucleic acids encoding bone marrow polypeptides, useful in diagnostic
PT	and gene therapy -
XX	
PS	Claim 10; Page 135; 392pp; English.
XX	
CC	AU14602-AU14794 represent novel bone marrow polypeptides of the
CC	invention. The proteins and corresponding coding sequences may be used
CC	in the prevention, diagnosis and treatment of diseases associated with
CC	inappropriate bone marrow polypeptide expression. For example, to treat
CC	disorders associated with decreased expression by rectifying mutations
CC	or deletions in a patient's genome that affect the activity of the
CC	polypeptides by expressing inactive proteins or to supplement the
CC	patient's own production of the polypeptide. Additionally, the nucleic
CC	acids may be used to produce the polypeptides, by inserting the nucleic
CC	acids into a host cell and culturing the cell to express the protein.
CC	The nucleic acid and its complementary sequences may also be used as DNA
CC	probes in diagnostic assays to detect and quantitate the presence of
CC	similar nucleic acid sequences in samples, and therefore which patients
CC	may be in need of restorative therapy. The proteins may also be used as
CC	antigens in the production of antibodies against bone marrow proteins
CC	and in assays to identify modulators of their expression and activity.
CC	The anti-bone marrow protein antibodies and antagonists may also be used
CC	to down regulate expression and activity. The antibodies may also be used
CC	as diagnostic agents for detecting the presence of the protein in samples
CC	(e.g. by enzyme linked immunosorbent assay (ELISA)). The proteins
CC	may be used to regulate haematopoiesis activity, and consequently in the
CC	treatment of myeloid or lymph cell disorders; in tissue regeneration,
CC	such as wound healing; as a nutritional supplement; and in treatment of
CC	immune disorders such as severe combined immunodeficiency (SCID).
XX	
SO	Sequence 138 AA;
QY	Query Match 9.7%; Score 706; DB 22; Length 138;
	Best Local Similarity 98.5%; Pred. No. 2,4e-55;
Matches 132; Conservative 0; Mismatches 2; Indels 0; Gaps 0;	
Db	169 SKRIIVLGHKHTRRITCGCMNAENLKALSGEDKMITYSNBEGTIRIQTVRSRFXNNQFL 228
	5 sKRIIVlghkhtrritcgcwnaenllalgsedkmitysnbgdlrtqtvrsrpsnmqfl 64
QY	KMDRTSAESMSISVLGKTLPFLNLNEPPNADLEQQOFGNIVCYNWYGDGGRIMIG 288
Db	65 mkmdrtsaesmsisvlgkttlfflnlnhepnadletqgdfgnlvcywygdgrlmig 124
QY	288 FSCGFHVVISITHTG 302
Db	125 fscgfhwvisithtg 138

RESULT 7
AAU33123
ID AAU33123 standard; Protein; 109 AA.
XX
AC AAU33123;
XX
DT 18-DEC-2001 (first entry)
XX
DE Novel human secreted protein #3614.
XX
KW Human; vaccination; gene therapy; nutritional supplement;
KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.
XX
OS Homo sapiens.
XX
PN WO200179449-A2.
XX
PD 25-OCT-2001.
XX
PF 16-APR-2001; 2001WO-US08656.
XX
PR 18-APR-2000; 2000US-0552929.
PR 26-JAN-2001; 2001US-0770160.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX
DR WPI: 2001-611725/70.
XX
PT Nucleic acids encoding a range of human polypeptides, useful in genetic
PT vaccination, testing and therapy -
XX
PS Claim 20; Page 710; 765pp; English.
XX
CC The invention relates to novel human secreted polypeptides. The
CC polypeptides and antibodies to the polypeptides are useful for
CC determining the presence of or predisposition to a disease associated
CC with altered levels of polypeptide. The polypeptides are also useful for
CC identifying agents (agonists and antagonists) that bind to them. Cells
CC expressing the proteins are useful for identifying a therapeutic agent
CC for use in treatment of a pathology related to aberrant expression or
CC physiological interactions of the polypeptide. Vectors comprising
CC the nucleic acids encoding the polypeptides and cells genetically
CC engineered to express them are also useful for producing the proteins.
CC The proteins are useful in genetic vaccination, testing and
CC therapy, and can be used as nutritional supplements. They may be used to
CC increase stem cell proliferation; to regulate haematopoiesis; and in
CC bone, cartilage, tendon and/or nerve tissue growth or regeneration;
CC immune suppression and/or stimulation; as anti-inflammatory agents; and
CC in treatment of leukaemias. AAU29510-AAU33304 represent the amino acid
CC sequences of novel human secreted proteins of the invention.
XX
SQ Sequence 109 AA;
XX
Query Match 4.1%; Score 298; DB 22; Length 109;
Best Local Similarity 59.3%; Pred. No. 1.1e-18;
Matches 64; Conservative 9; Mismatches 35; Indels 0; Gaps 0;

OY 962 KOMQSVIRIYDHLNNEKAVNIVRETOSLDGAKWVAREFLOJGYSALIOFLVMSKCN 1021
DB 2 EGPYSISISLIDGCPMPVILARXTAADESLVPTIMLGLADYGPVLYSLVIXVOLT 61
OY 1022 EAFITLAQGNKMEITYADITIGSEDTNEDYOSIALYFEGGKRLQAKGF 1069
DB 62 AATLAGNQSKMEIYADIGSEDTNEDYSLYALYFEREMYLGAQKFI 109
RESULT 8
AAM79130

ID AAM79130 standard; Protein; 1749 AA.
XX
AC AAM79130;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human protein SEQ ID NO 1792.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation.
XX
OS Homo sapiens.
XX
PN WO200157190-A2.
XX
PD 09-AUG-2001.
XX
PF 05-FEB-2001; 2001WO-US04098.
XX
PR 03-FEB-2000; 2000US-0496914.
PR 27-APR-2000; 2000US-0560875.
PR 20-JUN-2000; 2000US-0598075.
PR 19-JUL-2000; 2000US-0620325.
PR 01-SEP-2000; 2000US-0654936.
PR 15-SEP-2000; 2000US-0663561.
PR 20-OCT-2000; 2000US-0693325.
PR 30-NOV-2000; 2000US-0728422.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
PI Zhao QY, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AJ, Yang Y, Wehrman T, Goodrich R;
XX
DR WPI: 2001-476283/51.
DR N-PSDB; AAK52263.
XX
CC Nucleic acids encoding polypeptides with cytokine-like activities,
CC useful in diagnosis and gene therapy -
XX
PS Claim 20; Page 4152-4155; 6221pp; English.
XX
CC The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC encoded polypeptides (AAM78323-AAW80302) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.
CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
CC (AAM80020) are omitted as the relevant pages from the sequence listing
CC were missing at the time of publication.
XX
SQ Sequence 1749 AA;

OY 39 VEMKRIEFLKXTWLGAPIQ-FAMQKTSQNVLAVTGADYIKIFDRHGOKEINLPGNC 97
DB 1 MHLKHLITLSPQDGAQVTCMAWESGNAKF-avctvdvrvlllyehgerf----- 50
OY 98 VAMDMKDGQDVLAAVTAEKSSCIYLMWDANTNKTSQDNGMRDQM-SFLMSKXGSLFAYGT 156
DB 51 -----dkrtskpadmkxyrkksymvkxgmalfspdstkxtaiqg 85

```

QY 157 VKGNLXIV--NHOTSRKIPVIGK--HTKRITCGMNAENLXALG---GEDKMITVSNQEG 209
Db 86 tdnlllyvkkigedvqdkvicknfifqtsavtcqpwpaeylivlqglaegkvrlantknks 145
QY 210 DTRGTOVRESEPNMOMFLKMKDDRTSAESMISVVLGKKTLEFLNLNEPDNPALEFQO 269
Db 146 stlygt-----eeyvslctncsgkglaghadqtlvrf-----ddegsgsg 191
QY 270 DFGNIYC---YMWGDGRIMIGFSCGHFVYISTHTELGCEIFQARNHNDNLSTAVSQ 325
Db 192 klvnhpcppyalawatslvaag--cdrlvaygkeghmlqtdydrspqrefltsvss 249
QY 326 TLAKVATCCG--NCIKIDLDLDMQVYILNLDENK-----GLGTLMTDQGLLAL 376
Db 250 psgsgvnlgsydlrvfnwiprslv-----eakpkelnlytltalawkkdgsrlcv 303
QY 377 STGRSLHVEFLTKPLIGDASTRIAY-----LTSL--LEVTVANPYEGELPITVS--V 426
Db 304 gtlcgvvegf-----dcclrrslvknkfeltyvgpsqvivknlsqgrlvvllkshygy 355
QY 427 DVERNFYAVGLYLAVGMNNRANFYVIGENAVKKLKMEVLGYVA-----STCL-- 475
Db 356 eve-evklilgkerylvahtsct--lllgdlntrlselawqsgnqkyffenenvcmlf 412
QY 476 -----HSDYAALFEKGVLHLIESEILDQOE---ERETRLFPVADDCRILCH 521
Db 413 nageltlveymndlgsvtetfemphllsvrlnercqrgtdnkklaylidlktiaivd 472
QY 522 ALTSDFLIGTDGVVOYFIEDMOFVNDYRHPYSVKRIRPDNGRLVEIDKSDGFVY 581
Db 473 -----llggyllgltvshesgrvdlne-----tghkllfrdx----- 506
QY 582 CPVNDATYETRPDSPT--IGVLMENPMKGVFIAD-----DDKYTYTVFH 627
Db 507 --lrlhlydiescskmlnlfcaymwvpsdvlvagrnslcwuyalaeperutmlftr 564
QY 628 KDTI-----OGAK---VILASTKVPFAHKRPLLYNGELTQOTGSKVNNIYLSHGLSN 680
Db 565 gdtvlgderggsktevmwmevltvayt-----ldegllfgrtaldgn--ylyrataflet 617
QY 681 LKDXGPDELRLPMLANLMLKRFSDAMEMCRILNDEAAMNELLARACILNMEVEFAIRYVR 740
Db 618 l-----empr-----eteamwktlskialeargllhaercfsa 650
QY 741 IGVNGIVMSLEQIKGT-----EDYNLAGHLAMFTNDYNAQDLIYLAASSCPPIA 788
Db 651 lgvvakarflhetneladqvsreygsgtdfyvtrarlamlknyklaemlileqnaeve 710
QY 789 ALEMRRLQHMDSALQALAKHLAPDQIFPISKEVAYIQL-----EFAGDYVNAL 835
Db 711 amgmvygelhvwdeclavaeekyhpaleklrrsyqvwlmdeqgeeragelgesqdglaai 770
QY 836 AHYEKGTGDNKEHDEACLAGVAOMSTRMGDIRRGVNAQLKHSRYLKRDCGAILENMKO 895
Db 771 slvylk-----aglpaakaarl-----vltre--ellantel 798
QY 896 FSEAAQLYEKGLVYDKAASVYIRSKNMAKVGDLPHVSSPKIHQ-----YAKAE-A 947
Db 799 vehltaeilgelajera-----gdlfeklnhpqkalecyrknaifnakavela 845
QY 948 DGRYKEAVVAYENA-----KOMOSVIRIYLDHLNPREKAVNIVEETOSLDGAKK--V 997
Db 846 rlafpeevvnlleaewghlvqgqldaalnhyle-----arcsilaaleaajargawkk 898
QY 998 AREFLQGDYSAIOFLVMSKCNNEAFTLAQH-----NKMEIYADIGSEDTTNEQOSI 1053
Db 899 alylaldtdntasky-----yplvaqhyaslsqeylaeelytkgdrtkd--ai 945
QY 1054 ALYFEGEKRLQAGKFFLLC-----GOYSRALKNHLKPCSSSEDNV 1093
Db 946 dmytga--grwegahklamkcmrpedvsvlyltqgemekgkyreaerlyvtv--gepdl 1002

```

```

QY 1094 AIEK-----AIEVGOAKDELTLNQLDHL-LIS---ENDGMPKDAKYL----- 1133
Db 1003 aicmykxkhhlyddmrlrvghkhpdlisd-----tlhlgkelaegrldeeyhylaegewk 1059
QY 1134 ---RLYMALQYREAAQTAAIIAREEQSAGNVRNADVLSMYAELKSQRIKIPSEMATN 1190
Db 1060 atvmmyrasglweeay-----vartgga---nahkhvaylwak----- 1096
QY 1191 LMLHSIILVKIKHKNDDHKKGARMRLIRVANNNISKFPSSHVPIILTSTVIECHRAGLKNSA 1250
Db 1097 -----slggeaavrllinklg-----lleaavd-haa--dnscs 1125
QY 1251 FSPFAMLMREPEYRSK---IDAKYKKRTEGWRPRDISEIEEA 1289
Db 1126 fefafelstlalkhktpevhkyamfie-----degkfeea 1161

RESULT 9
ABB58828
ID ABB58828 standard; Protein; 1205 AA.
XX
AC ABB58828:
XX
XX 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 3276.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX
XX Drosophila melanogaster.
XX
XX W0200171042-A2.
XX
XX 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US09231.
XX
XX 23-MAR-2000; 2000US-191637P.
XX
XX 11-JUL-2000; 2000US-0614150.
XX
XX (PEKE ) PE CORP NY.
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
XX
XX WPI; 2001-656860/75.
XX
XX N-PDB; ABL02931.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions -
XX
XX Disclosure; SEQ ID NO 3276; 21pp + Sequence listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (ABL1840-ABL16175) and the encoded proteins
XX (ABB5737-ABB72072).
XX
XX The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pcl_sequences.
XX
XX Sequence 1205 AA:

```

Query Match 4.08; Score 289; DB 22; Length 1205;
 Best Local Similarity 17.9%; Pred. No. 5.8e-16;
 Matches 261; Conservative 219; Mismatches 530; Indels 452; Gaps 62;

QY	60	AMQRTSNGWYLVAGDADYVTKF-----DRIGQKRSEINLPGNVANDMDKODGVAVIAE	114
Dd	21	awnkeeg-yiaveatcdgllkvikldgatprgskgylaavsn-lsmqclddh-----k	72
QY	115	KSSCYLW-DANTNKTSQLDNGM-----RDMSEFLMSKVS---	150
Dd	73	esvrvvtvndqgkrltsddcdgvlmvwmllkgsyeemthdrkkslvsaamswtsdgsrlc	132
QY	151	-----FLAVGTVGNLKIYNHQTSRKIPVLGKHTK--RITGCGMAEN--LXALGED	199
Dd	133	lyvedgalivgsvdgn-----rifgelkgthltgygwepdnrlilfalaenge	180
QY	200	-----KRIIV--SNEBGTI-----KQIVSESPXNMQFELMKMD	232
Dd	181	chlYdngqnfamkllhdcvnlsgsssrgrhtaslscwfsgrvvgstkrp-----	229
QY	233	DRTSAEISMISVVLGKKTFLFLNLNEPDNPADLEFQODFENICYCWMYDGRIMIGFSG	292
Dd	230	-----viacYengrvqimrendapa--ifdtgmrvda-kwhdgvla--icg	276
QY	293	HFV--VISTHTGELGOELFOGRANK-----DNLTSIAVSQTLNKVATGDNCKIKIO	341
Dd	277	tltaavptsgdrtnqvcfyfsglkiyrltkvpgtdltsiswegkelrlamavdsfily--	334
QY	342	DLVLDKDMYVILNLDEEKKGIGTISWTDQCLALAS--TORGSLHVFYTLRP-----	391
Dd	335	-fanltpdyicwyfeketvflnsgsalrespmvslftmlvtnsqslfikeveptlclases	393
QY	392	---ILGDGC-STRIAYLTL-SLEEVYTNAPPE-----GELPRIVS--ADVBNPFA	434
Dd	394	ehcvlgyavcvsnnkelatlstle-nrsnpaddrvyqlllensigtvdsdkytdlprctfyg	452
QY	435	VGLVHNLAVGMNRR--AMFYVLGENA-----VKRLKDMEYLGTVAISICLSDYAALFEK	487
Dd	453	insyvalaasegelllwhynhrpkasellhnvkarkefrf-----hldtprgvema	503
QY	488	VOLHLISEIILDAOEBERTLFRPVDKCKILCHALTSPDLITGDTGVGYFYIEDMOF	547
Dd	504	kdlmlssssgqghstqr-----glsdp--lcalatseklllvaresgaingelsanval	555
QY	548	VNDYHNPRYSVKRPPDPNGRTFLYRIDKSDGFVCPVNPADTYEIRPEFPTIGVLWE--NM	606
Dd	556	rnrhlmakyykmatncnstaaildmng--vmclldldnretqlnfrstverkdvawavsw	614
QY	607	PMDKGVFIAYDDDKVYTVVFEHKTIOGAKVILLAGSTKVPFAHKRPLLNLNGELTCTOSGK	666
Dd	615	atdnprlllal-mektrmlyifrgndpe-epscsg-----ylctfedle	655
QY	667	VNNIYLS---THGFLSNLKDXGPDELPRMLAH--NMLKRFSDAMCMCLINDEBA----	717
Dd	656	itsvlllddlslsvgelqn-----fshliqrlvkslrtdldllenhvledakqfi	703
QY	718	-----WNELARCAJHMEVEFAIRVYRIGVGIVMSLEQIKIEDYVNLGHLMLF	769
Dd	704	ednphrptlwrlllaealkklelecaenatrfycalpygiklvtlrllnshkelgraelstf	763
QY	770	TNDYNLADOLYLIASSCPALAEEMRRDLOHDSALOLAK-----HL	809
Dd	764	ygefeeeaeeklylddcdrrldalrelmtclcdwfrvvqlyimgsgvsdqmeiawrelghf	823
QY	810	APDQIPFISKEYAIQLEFAGDYVALAHYKEGITTGDNKEHDEACLAGVAMQMSIRMGDIRR	869
Dd	824	anlrswareareyekshnylgegymealylhe-----qtdldlek	860
QY	870	GVNQALKHPRSVYLNKRDGATILENNKQFSEEAOLYERKGLYYDKAAV--YIRSKNNAKGD	927
Dd	861	cv-erlpekspdlrk-laemlaasygmseseavghlr--fgdqkaavacvnlrtprgeave	916
QY	928	LLPHVSSKILHLOZAKAKE--ADGRKMAVVAVYENAKAQMSQVIRYLDHLNNPEKAVNI	984
Dd	917	laqrfqprpqvqtlakhaaqqlgegrlkeaeliemgrna-----	953
QY	985	VREQTSLDGAAMVAFRLQLODVGSAIQFLVMSKCNNEAFTLQAQHNKMEIYADIGSED	1044

Dd	954	---	grhl	aarllsq-	maeregekrapl	lrkiklyv	laallaehuk-----ava	999
Oy	1045	TNTEYOSIALYFESEKRYLQAGKFELLGCGOYSRALKHFLCPSSEDVAIEMAIETVGQ	1104		:	:	:	:
Dd	1000	tteidyas-----grrtll-----dsiale-----	1019		:	:	:	:
Oy	1105	AKDELLTNOLIDHLGENDGMKPDAKKYLFRLYMALKOYREAAQTALIIAREEOSAGNYRN	1164		:	:	:	:
Dd	1020	-----daaaierlwhcaeyh-----fmllaqqlrfg----	1047		:	:	:	:
Oy	1165	AHDVLFMSWAEELKSOKIRIPSEMATNTMLTISYLLVKTHVKNNGDHMKGARLRIVANNIS	1224		:	:	:	:
Dd	1048	-----lvhsavvtavrllldyedvip-----	1067		:	:	:	:
Oy	1225	KFPSHIVEPLTWSTVIECHRAGLKNSAFSFAAMLMPREVRSKID---AKYKKRIEGWRR	1280		:	:	:	:
Dd	1068	--pehiysallalaasacdratfctskaf-----mklegahlpcaatlqryeelaagflak	1119		:	:	:	:
Oy	1281	PDISIEEATTPPCPFCKFLBCECELLCPGCNKNSIPYCITAGRHMUK--DWM--TVCPHC	1335		:	:	:	:
Dd	1120	ydpeditgdvdycdygcgyvpdpspspcpecnarfpacisgsqklptgpnaiwtictchc	1179		:	:	:	:
Oy	1336	DFPALYSELKIMLNTESTCPMC	1357		:	:	:	:
Dd	1180	aapt-----elshrtpclpic	1194		:	:	:	:
 RESULT 10								
ID	AAU33013							
ID	AAU33013	standard; Protein: 72 AA.						
AC	AAU33013;							
XX	18-DEC-2001	(first entry)						
DT								
XX								
DE	Novel human secreted protein #3504.							
XX								
KM	Human; vaccination; gene therapy; nutritional supplement;							
KW	stem cell proliferation; haematopoiesis; nerve tissue regeneration;							
KM	immune suppression; immune stimulation; anti-inflammatory; leukaemia.							
XX								
OS	Homo sapiens.							
XX								
PN	WO200179449-A2.							
XX								
PD	25-OCT-2001.							
XX								
PF	16-APR-2001; 2001WO-US08656.							
XX								
PR	18-APR-2000; 2000US-0552929.							
PR	26-JAN-2001; 2001US-0770160.							
XX								
PA	(HYSE-) HYSEQ INC.							
XX								
PI	Tang YT, Liu C, Drmanac RT;							
XX								
DR	WPI; 2001-611725/70.							
XX								
PT	Nucleic acids encoding a range of human polypeptides, useful in genetic							
XX	vaccination, testing and therapy -							
PS	Claim 20; Page 698; 765pp; English.							
XX								
CC	The invention relates to novel human secreted polypeptides. The							
CC	polypeptides and antibodies to the polypeptides are useful for							
CC	determining the presence of or predisposition to a disease associated							
CC	with altered levels of polypeptide. The polypeptides are also useful for							
CC	identifying agents (agonists and antagonists) that bind to them. Cells							
CC	expressing the proteins are useful for identifying a therapeutic agent							
CC	for use in treatment of a pathology related to aberrant expression or							
CC	physiological interactions of the polypeptide. Vectors comprising							
CC	the nucleic acids encoding the polypeptides and cells genetically							

CC engineered to express them are also useful for producing the proteins.
CC The proteins are useful in genetic vaccination, testing and
CC therapy, and can be used as nutritional supplements. They may be used to
CC increase stem cell proliferation; to regulate haematopoiesis; and in
CC bone, cartilage, tendon and/or nerve tissue growth or regeneration;
CC immune suppression and/or stimulation; as anti-inflammatory agents; and
CC in treatment of leukaemias. AA029510-AA03304 represent the amino acid
CC sequences of novel human secreted proteins of the invention.
XX
SQ Sequence 72 AA:

Query Match 3.8%; Score 274.5; DB 22; Length 72;
Best Local Similarity 85.7%; Pred. No. 7e-17;
Matches 54; Conservative 3; Mismatches 5; Indels 1; Gaps 1;

QY 274 IVCYMWGDRIMGFCGHPVYISTHTGELGGEIFQARHKNDLJISAVSQTINKATC 333
Db 9 ILCYR-YgdgrImfgscghxvsthtgclgclfgarhkdltlsavsqtlkvatc 67

QY 334 GDN 336
Db 68 em 70

RESULT 11
AAM79062
ID AAM79062 standard; Protein; 1241 AA.
XX
AC AAM79062;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human protein SEQ ID NO 1724.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation.
XX
OS Homo sapiens.
XX
PN WO200157190-A2.
XX
PD 09-AUG-2001.
XX
PE 05-FEB-2001; 2001MO-US04098.
XX
PR 03-FEB-2000; 2000US-0496914.
PR 27-APR-2000; 2000US-0560875.
PR 20-JUN-2000; 2000US-0598075.
PR 19-JUL-2000; 2000US-0620325.
PR 01-SEP-2000; 2000US-0654936.
PR 15-SEP-2000; 2000US-0663561.
PR 20-OCT-2000; 2000US-0693325.
PR 30-NOV-2000; 2000US-0728422.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
PI Zhao QA, Wang J, Zhang J, Ren F, Chen R, Wang ZM;
PI Xue AJ, Yang Y, Wejrtman T, Goodrich R;
XX
DR WPI: 2001-476283/51.
DR N-PSDB: AAK52195.
XX
PT Nucleic acids encoding polypeptides with cytokine-like activities,
PT useful in diagnosis and gene therapy -
XX
PS Claim 20; Page 4060-4063; 6221pp; English.
XX
CC The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC encoded polypeptides (AAM78323-AAK80302) that exhibit activity relating to

CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activity/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.
CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
CC (AAM80020) are omitted as the relevant pages from the sequence listing
CC were missing at the time of publication.
XX
SQ Sequence 1241 AA:

Query Match 3.6%; Score 263; DB 22; Length 1241;
Best Local Similarity 18.2%; Pred. No. 1.4e-13;
Matches 245; Conservative 197; Mismatches 506; Indels 396; Gaps 54;

QY 116 SSCYIWMANTNNTSOLDNNGMDQMSFLWSKVSFLAVGTGKGLXLYNHQTSRKIPV 174
Db 113 ssdglwspqkvsyk-ihssskllccswndgqyalgmfgllstlnkgeekvk 170

QY 175 -LGKHTKRITCGW-----NAENLXALGEDKMT-----TVSNOEG 209
Db 171 rpggsislpiwscmpsrstwestfmmarenda---edvlnrlylqelprslkaavysgq 227

QY 210 DTRQTOVRSEPKXNMOFLKMDRTSAESMISV-LGKKTLEF-LNLERNPADLEF 267
Db 228 sgeeeeeeedspr-----ddnleerndllavdgqkvsfygskqigkralnf 281

QY 268 QDPEGNIVCYMWGDR-IMGFCGHPVYISTHTGELGGEIFQARHKNDLJISAVSQT 326
Db 282 dp-----ccisyltkgeyllg--gsdkqvsft-----kqgv----- 312

QY 327 LNKVATCGDNCIKIQDVLDEKDMIVLNLDEENKGLTSLWTDGQDLASTORGSLNHF 386
Db 313 --flgtvg-----eqnswvltcgakrdsnyvvvgcqdgtfsf 348

QY 387 LTKLPILGDACSTRIVATLSLEYTVANPYEGELPTVSVDENFVAVGLNHLAVGMN 446
Db 349 qlfstvhgylkdrayrdmtdvghlft-ekvnrkckelvjkkaiaylnlaqlpe 407

QY 447 RAMEYVGENAVKVKKMDMETLGTVASICLSHDAALFEKGVOLNLESETLDAEER-E 505
Db 408 kiliyels---edlsdmhy-----rvkekllkfecnl-lvvaanhllcoqekrlq 455

QY 506 TRLEPAVDK-----CRILCHALTSDFLYGTDTGVVOYFYEDMOFVNDYRNPV 555
Db 456 clsfsgykerewmesllyrlykviqppregllvglnkgqllkifvdlfaiylkqat 515

QY 556 SVKKIFPPDPKTRLVFIDKSDGDFVPCPVNDATYEIPDFSTIKGVLMEMPRDKGYFT 614
Db 516 avrcldmsarikklavdendctlv---dlctkellfgeprnsvawmncqcedmclsfsg 572

QY 615 -AYDDDKVYTVFHKDTIGCAKVTLAGSTKVPRAHKRLLLYNG-ELTCQPGOSGVNNIYL 672
Db 573 ggylnlkastrfyrnqklqg---fvvg-----yngsklfc----- 604

QY 673 STHGFSNLKDXGDELRPMLANHLMLKRFSDAMEMCRILNDEAMNMLARACLNHMEVE 732
Db 605 -lhvf--sisavevpgsamygy-ldtkllkeayqalcjvtdvdwelmleaegldfe 660

QY 733 FAIRVYRRIG--VGIVMSLEQIK--GIEDYVNLACHLAFMTNDVYLAODYLASSCPT 787
Db 661 takkafirvgdlylleiissieerkkrgetnndlladvfsyggkfkneaklykrsgen 720

QY 788 AALEMRDLOHMSALDGLANHLARPDQIRFISKEVAIDLEFAGDVNMLAYEKGITGDK 847
Db 721 lalemytdl-----cmfeyakdflyg-----sgdpx 745

QY 848 EHDEACLAGVAOMSIRMGDIRGVNOLAKHPSRYLKRDCALLENMKOFSEAOLEYEGL 907

```

Db 746 e-----tkmltkqadwar-----nlikepkaavemyisag 775
QY 908 YKDKASVYIRSKNMKAGDGLPHVSSPKIHLOYAKAKKADGKRYKAAVAYENAKMOSY 967
Db 776 ehvkalet-cgdhgvw--dm1ldl-----arkldaeiepl1-----1 810
QY 968 IRIYDLNPNPEKAVNIVRETOSLDGAKMVAFFLOLDGYSALQPLVMSKCNNEAFTLA 1027
Db 811 catylkkldspya-----aetylkmgdlkslvqlhvetqrwdeafalg 854
QY 1028 QQHNMKEIYADITIGSEDTTNEYOSIALYFEGEKRYLQAGKFFLLCGQYSRALKHFLKCP 1087
Db 855 ekhpelk-----ddlympyagwlaendrfeeqakthkagrqreav----- 895
QY 1088 SSEDNVAIMEATFVGQANDELLTNOLIDLGENGMKDKAYFLRLMALKQYREAO 1147
Db 896 -----qvleqltnavaesrf-----ndaayy--wmlsmqclidiq 930
QY 1148 TAIIRBOSAGNRYNADHVFESMYAEELKSOKIKIPSEMATMLHSHYLVKIHVKNG 1207
Db 931 dp---aqktumlgkfyh-----fgriael-----yghyhalhrhtedp 965
QY 1208 DHMKGBMLIRVANNISKPSHVPILITSTVIECHR---AGLKNASFSAAMLMPREYS 1264
Db 966 fsvhrpelk---nistrllhslpdxpjskxklltflakqskalyarlarhaydk 1021
QY 1265 ---KIDAKYKKKIE---GMVRRPDISEIEEATTPCPFCFKFLPEELLCPGCKNSIPYC 1317
Db 1022 lqglyparfqsietglttlrakphndee-----lvplc-----yr 1059
QY 1318 IATGRMLKDMVTVCPCHPALYS-----ELKIMLNTSESTCPM 1356
Db 1060 cslmpnlhlnhgvncincqpfifssasydvhlhvefyldgealtdlealslidlvlrpk 1119
QY 1357 CSFRLMAOLKRTSDCTOYLREE 1380
Db 1120 rddr---qlelannssqllrlve 1139

RESULT 12
AAB43197
ID AAB43197 standard; Protein; 1252 AA.
AC AAB43197;
DT 08-FEB-2001 (first entry)
DE Human ORFX ORF2961 polypeptide sequence SEQ ID NO:5922.
XX
XX Human; open reading frame; ORFX; detection; cytosolic; hepatotropic;
XX vulerary; antiproliferative; antiparkinsonian; neurotrophic; neuroprotective;
XX anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
XX immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
XX hypotensive; dermatological; immunosuppressive; antinflammatory;
XX antiviral; antibacterial; antifungal; antineumatic; antihypoid;
XX antianemic; gene therapy; cancer; proliferative disorder; hypertension;
XX neurodegenerative disorder; osteoarthritis; graft vs host disease;
XX cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
XX cholesterol ester storage; systemic lupus erythematosus; infection;
XX severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
XX allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
XX bone damage; cartilage damage; antinflammatory disease; coagulation;
XX thrombosis; contraceptive.
OS Homo sapiens.
PN WO200058473-A2.
XX
XX 05-OCT-2000.
XX
XX 31-MAR-2000; 2000WO-US08621.

```

```

PR 31-MAR-1999; 99US-0127607.
PR 02-APR-1999; 99US-0127636.
PR 05-APR-1999; 99US-0127728.
PR 30-MAR-2000; 2000US-0540763.
XX
XX (CUBA-) CUBAGEN CORP.
XX
XX Shinkets RA, Leach M;
XX
XX WPI: 2000-602362/57.
XX N-PSDB; AAC77406.
XX
XX Novel nucleic acids and peptides derived from open reading frame X,
XX useful for treating e.g. cancers, proliferative disorders,
XX neurodegenerative disorders and cardiovascular disease -
XX
XX Claim 11; Page 5103-5106; 5507pp; English.
XX
XX AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
XX which represent the human ORFX open reading frames 1 to 3161. The ORFX
XX sequences have activities such as: cytosolic; hepatotropic; vulerary;
XX antiproliferative; antiparkinsonian; neurotrophic; neuroprotective;
XX osteopathic; anticonvulsant; antarthritic; immunosuppressant;
XX immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
XX antidiabetic; hypotensive; dermatological; immunosuppressive;
XX antinflammatory; antibacterial; antiviral; antifungal; antineumatic;
XX antihypoid; and antianemic. The sequences can be used for determining
XX the presence of or predisposition to, or preventing or treating
XX pathological conditions associated with an ORFX-associated disorder. The
XX nucleic acids can be used to express ORFX proteins in gene therapy
XX vectors. The proteins and nucleic acids may be used to treat cancers,
XX proliferative disorders, neurodegenerative disorders, osteoarthritis,
XX graft vs host disease, cardiovascular disease, diabetes mellitus,
XX hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
XX erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
XX bacterial or fungal infection, malaria, autoimmune disorders, asthma,
XX allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
XX nocturnal haemoglobinuria, antinflammatory disease; to enhance
XX coagulation; to inhibit thrombosis; and as a contraceptive.
XX
XX Sequence 1252 AA;

Query Match 3.6%; Score 260.5; DB 21; Length 1252;
Best Local Similarity 18.2%; Pred. No. 2.3e-13;
Matches 245; Conservative 197; Mismatches 507; Indels 395; Gaps 54;

QY 116 SSCITLMOANTKTKTSOLDNGMDQMSFLMSKVSFLANGTYKGNLXIYHQTSKRIPV- 174
Db 71 ssdfglwspqekvsak--hkssskllccswtndgqylagmfnlgllstrnkgeekvkie 128
QY 175 -LGKHTKTRTCGCM-----NAENLXALGEDKKMI-----TYSNOEG 209
Db 129 rpggsfslpwsicwpsrswesfmmnredea---edvlnvnylgelpslksavysgq 185
QY 210 DTRQTVQVRSPEXNMQFLMKMDRTSAESKITSVY-IGKTKLFF-LNLNEPDNPADLEF 267
Db 186 saaeesepeeeddpr-----dnlleerndllavadgqkvsfyglsgkqglkavralnf 239
QY 268 QQDFGIVCYMNYGGR-IMIGFSGHFVYISTHNGELGQELFQARNKNDNTSLAVSQT 326
Db 240 dp-----ccisylftkgey11lg--gsdkqvs1ft-----kdyv----- 270
QY 327 LNKVATCGDNCKIKIDVLDMKMYVLINLDEENKGLGTLSPWTDGQLALSTORGS1HF 386
Db 271 --rlgtvg-----egnswwtcqakpsdnyvvvgcgqglisfy 306
QY 387 LTKLPILGDACSTRITAVYLSLEVTVAANPEGELPITYSVDEVENPVAAGVLAHVGANN 446
Db 307 qlifstvnlglykdraydsmtdivgnhlt-egkvrlkckelvkklaiyrnlaigipe 365
QY 447 RAEFVYLGEMNAVKIKDKMEYLGTVASICLHSDYAAALPEGKQVOLLISEILLDAQEE-R 505

```

```

Db 366 klllyels---edlsdmhy-----rvkekllkfeonl-lvvcanhllqcekrld 413
OY 506 TRLEPRAVDK-----CRILCHALTSDFLIYGTDRGVGVYFIEDMQOFNDVRHHP 555
Db 414 clsfsvkeewmealllyikvlgppgrrgllylknqgllkifvdhlfavllkqat 473
OY 556 SVKKIRPPDNGTRLVPEIDKSDGFVVCVPVNDATVEIPDPSPIKGYLWENWMDKGVFI- 614
Db 474 avrcldmasarkklavdendctly---didckellfgepanasvamtqcdmclcfsg 530
OY 615 -AYDDKVVTVYFHKDTIGACAKYILAGSTKVPRAHKRPRLLYNG-ELTCTOSGKVNNTYL 672
Db 531 ggylnlkaestfvrhqrklqg---fvvg-----yngsklfc----- 562
OY 673 STFGFSLNKDKGPDRLPRLAHNLMLKRPDAMKCRILINDAANNELARACTHHMEVE 732
Db 563 -lhvf--slsvevpsapmyqy-lqrklfkeayqiaclgvtdctdvrclamealegldfe 618
OY 733 FAIRVYRRIGN---VGIVNSLEQIK--GIEDVNLGLAGHLAMFTNDYNLAQDLYLASSCPT 787
Db 619 takkatfiryqdllylellsleerkkrrgetnndlfladvfsyqgkfhcaaklyktrsgnen 678
OY 788 AALEMRDLOHMDSALOQLAKHLAPDOIPTISKRYAIQLEPAGDYVNAALAHYEKGITGDNK 847
Db 679 lalemvtdl-----cmfeyakdfllg-----sgdprk 703
OY 848 EHDEACLAGVAQMSIRMGDIRRGVNOALKNPSVLKRDGAILENKKESEAOLEKGL 907
Db 704 e-----tkmltlkqadar-----nlkepkaaemylsag 733
OY 908 YYDKAAVYIRSKNMKAVGDLPRHVSFKIHLQYAKAKKADGRYKAVVAUYENAKOMQSV 967
Db 734 evkatelei-cgdhgwv---dmldldi-----arkdkaeerpll-----l 768
OY 968 IRIYDLHLNPEKAVNIVRETOSLGCARKVAFRLQLDGYSAIQFLVNSKKNNEAPTLA 1027
Db 769 catylklklspgya-----aeyllkmgdlkslyvqlhveqqrwdeafalg 812
OY 1028 OOHNKKEITADIIGSEDTNEDYQSIALYFEGSEKRYLQAGKFFLGGVSRALKHFLKCP 1087
Db 813 ekipeik-----ddlymryqvlaendrfeeaqkafhkagrrgreav----- 853
OY 1088 SSEDNVAIEMAIETVGOAKDELITNQLIDHLGENDGMPKDAKYLRFLYVWALQYBEAQ 1147
Db 854 -----qvlleqlunavaesrf-----ndaayyy--wmlsmqclldiaq 888
OY 1148 TAIIRREOSAGNYRNANDVLFMTAEKLSQKIKIPSEKATMLMLHSTIILYKIHVKNQ 1207
Db 889 ad--paqcdmlygkfyh-----fgtrialael-----yhgylhalhrhdedp 924
OY 1208 DHKKGARMLIRVANNTSKRPSHIVPLITSVIECHR--AGLKNSAFSPFAMLMRPEYRS 1264
Db 925 fsvhrpccflf---nlsrllshlpkdcpsgjskvkllflaqsakalagryrlarhaayk 960
OY 1265 ----KIDAKYKKRIE--GMYRRPDISLEIEATTPCPFCFLLPCEELLCPGCKNSIPYC 1317
Db 981 lrglyparfqkelselqgltrikrpfhdeee-----lpylc-----yr 1018
OY 1318 IATGRHMLKDMTVCPCPCFPALYS-----ELKTMLTSTECPM 1356
Db 1019 csnfnpllnlmgvncncrpfllfssasydvlnhvefyaleegltdeeaialldlevlrpk 1078
OY 1357 CSERLNAOALKRISDCTOYLRTTEE 1380
Db 1079 rddr-----qlelannssqllrlfve 1098

```

```

DT 08-FEB-2001 (first entry)
XX
DE Human ORFX ORF2012 polypeptide sequence SEQ ID NO:4024.
XX
KW Human: open reading frame; ORFX: detection; cytostatic; hepatotropic;
KW vulnerability; antiproliferative; antiparkinsonian; neurotropic; neuroprotective;
KW anticonvulsant; osteopathic; antilarthritic; immunosuppressant; cardiant;
KW immunostimulant; thrombolytic; coagulant; vasotrophic; antidiabetic;
KW hypotensive; dermatological; immunosuppressive; antihypertensive;
KW antilarthritic; antifungal; antirheumatic; antihypertensive;
KW antianemic; gene therapy; cancer; proliferative disorder; hypertension;
KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW cholesterol ester storage; systemic lupus erythematosus; infection;
KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW bone damage; cartilage damage; antinflammatory disease; coagulation;
KW thrombosis; contraceptive.
XX
XX Homo sapiens.
XX
XX WO200058473-A2.
XX
XX 05-OCT-2000.
XX
XX 31-MAR-2000; 2000WO-US08621.
XX
XX 31-MAR-1999; 99US-0127607.
XX
XX 02-APR-1999; 99US-0127636.
XX
XX 05-APR-1999; 99US-0127728.
XX
XX 30-MAR-2000; 2000US-0540763.
XX
XX (CURA-) CURAGEN CORP.
XX
XX Shinketsu RA, Leach M;
XX
XX WPI: 2000-602362/57.
XX
XX N-PSDB; AAC76457.
XX
XX Novel nucleic acids and peptides derived from open reading frame X,
XX useful for treating e.g. cancers, proliferative disorders,
XX neurodegenerative disorders and cardiovascular disease -
XX
XX Claim 11: Page 3208-3212; 5507pp; English.
XX
XX AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
XX which represent the human ORFX open reading frames 1 to 3161. The ORFX
XX sequences have activities such as: cytostatic; hepatotropic; vulnerability;
XX antiproliferative; antiparkinsonian; neurotropic; neuroprotective;
XX osteopathic; anticonvulsant; antilarthritic; immunosuppressant;
XX immunostimulant; cardiant; thrombolytic; coagulant; vasotrophic;
XX antidiabetic; hypotensive; dermatological; immunosuppressive;
XX antinflammatory; antianemic; antiviral; antifungal; antirheumatic;
XX antihypertensive; antidiabetic. The sequences can be used for determining
XX the presence of or predisposition to, or preventing or treating
XX pathological conditions associated with an ORFX-associated disorder. The
XX nucleic acids can be used to express ORFX proteins in gene therapy
XX vectors. The proteins and nucleic acids may be used to treat cancers,
XX proliferative disorders, neurodegenerative disorders, osteoarthritis,
XX graft vs host disease, cardiovascular disease, diabetes mellitus,
XX hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
XX erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
XX bacterial or fungal infection, malaria, autoimmune disorders, asthma,
XX allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
XX nocturnal haemoglobinuria, antinflammatory disease; to enhance
XX coagulation; to inhibit thrombosis; and as a contraceptive.
XX
XX Sequence 1690 AA:
SQ

```

Query Match 3.6%; Score 259.5; DB 21; Length 1690;
 Best Local Similarity 19.0%; Pred. No. 4.9e-13;
 Matches 261; Conservative 187; Mismatches 463; Indels 463; Gaps 63;

```

QY 76 YVYIFDRHGQRSEIN---LPGNCVAMDMKDGDLAVIAEKSSCIYLMANTNKTSQL 132
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 25 yyykigedwgdkvcknckfjtsavetclqmpaeilylvfglaegk--vrlancktnksstl 82
QY 133 DNGMRQMSFLLMSKVGSLAVGTAVGNLXINYHOTSRIIPVLGK---HTKRITGCCWMA 189
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 83 -ygteesvavltncsgkqllsgnadqlivrylfddegsgesgkvlvncpppyalaw-a 140
QY 190 ENLXALGEDKMITVNSOGDTRROTQVSEPNKMOFLMKMDPRISAESM--ISVVG 247
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 141 tnsivaagcdrtlvaygkqgmqlqtdfysrdgeref-----ttavspbgsgsvvlg 192
QY 248 -----KTLFELNLEPNRADLEFQODFGNIVCYNWGDG--RIMIGSCGHF 294
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 193 sydlrlvfwlprsrw-----eakrkei--tnlyltltaawkrdgsrllcvgtlcg-- 242
QY 295 VVISTHTGLGGEIIPA-----RNHKDLTSLAVSOTLTKAVATCGDNCI-----KIQ 341
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 243 -----gveqfdcccltrsllykfkfeltygspqvavknlsgqlrvvlkshygyeve 292
QY 342 DLVDL-KDMVYILNDEENKGLTSLTSMDDGQLLALSTQSGSLHVFUTKLPIIGDACSTR 400
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 293 evklgkykelylvahkse-----tl-----llgdlntnr 320
QY 401 IAYLTSLEVTYVANPVEGELPTTVSVDBPNFVAVGLYHLAVGMNRWF-----YV 452
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 321 lseia-----wggs-----gnekylffenavcmifn 347
QY 453 LGEENAVKKIKLDMEXLTGTVASICLHSDYAAALFEGKQVLLIESELIDAE---ERETRLP 509
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 348 agetlvevgndtlsgsvtelfmp-----hlsvrlnerqrgvlednkklia 394
QY 510 PAVDDKCRILCHALYDFTLTGTDTGVQFYEDMQFNDYRNHPYSVKKIFRPDNGTLY 569
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 395 ylidktktiaivd-----lsgynlgvtvhesvtdvdelelne-----tghl 434
QY 570 VIDEKSDGFVYCVNDATYETIPDESPT--IKGLWENKPMDKGFIAVD----- 617
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 435 lfrdr-----lrlhlydiescsktmlnlfcsymgvpsdvlvagrnslcwmynl 486
QY 618 --DDKYYTVFPHKDTI---QGAK--VLAGSTKVPFAHKRLLLYNGELTQOTOSGKNV 668
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 487 eapervtmftlrgdvlglerggsktevmwmevgtlvayt-----ldegllfgrtaiddgn 541
QY 669 NYLSTHGFSLNLDKXDPDELPRMLAHNLMLKRFSDAWEMCRILNDEAMNELARACLNH 728
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 542 --yiratafletl-----emtr-----eteamwklckslalea 572
QY 729 MEYFAIRYVRRKIGNVGIYMSLEQIKGT-----EDYNLAGHLAMFTYNDYNLA 776
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 573 rqlhaercfsalgyvakarfjhetneladvrsreygsgeitdfygyrarlamleknykia 632
QY 777 ODLYLASSCPRIALEMRRDLOHMDSALOLAKHLAPQIPIFISKEYAYIOL----- 825
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 633 emfiflegnavveeamgmgyelhrwdecivavaeakghpaleklttrsyqwlmdrqgeerage 692
QY 826 --EFAGDYNALAHYKEGITGDNKEHDEACLAGVAQMSIRMGDIRRGVQALNHPSRVVK 883
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 693 lqesqgdgaiaaislylk-----aglpakaarl-----vlt 722
QY 884 RCGGAILENMKKOPSEAAQLYKGLYDKAASYINSKNNAKVGDLPHVSSPKHILQ--- 940
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 723 re--ellantelvehntaailkqelyera-----gdlfeklhnqkalecyr 767
QY 941 ----YAKAKE-ADGRYKKEVNAVYENA-----KOMOSVIRIYIDLHNLNPKAVNIVRE 987
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 768 kynaImkaealarlarlprvevnlleewgshlvqkylidaahlyle-----arcsika 820
QY 988 TQSLDGAAM--VAPFLOLDGYGSAIQFLVMSKCNNEAFTLAQOH---NKMEIYADITG 1041
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 821 leaalgarqwkakajylldqdrntasky-----yplvaghyaslqeyelaeeelyt 870

```

```

QY 1042 SEDTTNEDYQSIATLYFEGEKRYLQAKFFLLC-----GOYSRAIK 1081
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 871 kgdrtkd---aidmvtga-grwegahklamckmrepdevsylvltqagemekgkyreer 926
QY 1082 HFLKCPSSSEDNVAIEM-----AIEFVGAOKDELLTNOLDLH-LG---ENDGMPRD 1128
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 927 lyvtv--qepdlaftmkykhhklyddmirlyvghpdlisd---tvlhlgkeleaaqrlyqe 981
QY 1129 AKYLF-----RLYMALQYREAAQTALIIIRREOSAGNRYNRAHDVLSMYAELKS 1178
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 982 aeyhylqagewkatvnymyasglweeayr-----vartgga---nahkhvaylvak--- 1030
QY 1179 QKIKIPSEMATMLILHSYLLVKIHVKNGBHMKKAMRLIRVANNISKEPSHIVPLTSTV 1238
Db 1031 -----sllggaavrllnklyg-----llaa 1050
QY 1239 IECHRAGLKNASAPFAMLMREYRSK---IDAKYKKRKEGWRPRDISETEBA 1269
Db 1051 vd-haa--dnscfetafelstrlalxkhtpevnlkyamfle-----degkfeea 1095

```

```

RESULT 14
AAE01339
ID AAE01339 standard: protein, 993 AA.
XX
AC AAE01339;
XX
DT 17-JUL-2001 (first entry)
XX
DE Human gene 24 encoded secreted protein fragment, SEQ ID NO:204.
XX
KW Human; secreted protein; proliferative disorder; cancer; tumour;
KW foetal abnormality; developmental abnormality; haematopoietic disorder;
KW immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
KW inflammation; allergy; neurological disorder; Alzheimer's disease;
KW Parkinson's disease; cognitive disorder; schizophrenia; asthma;
KW skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;
KW cardiovascular disorder; angiogenic disorder; kidney disorder;
KW gastrointestinal disorder; pregnancy-related disorder;
KW endocrine disorder; infection; wound healing; vulnerability;
KW cell culture; chemotaxis; food additive; gene therapy;
KW binding partner identification.
XX
OS Homo sapiens.
XX
PN WO200134769-A2.
XX
PD 17-MAY-2001.
XX
PF 01-NOV-2000; 2000WO-US30040.
XX
PR 05-NOV-1999; 99US-0163580.
PR 30-JUN-2000; 2000US-0215130.
XX
PA (HGMA-) HUMAN GENOME SCI INC.
XX
PI Ruben SM, Komatsoulis GA, Wei P, Baker KP, Fiscella M;
XX
DR WPI; 2001-308781/32.
XX
PT New isolated nucleic acid molecule encoding a human secreted protein is
XX used in preventing, treating or ameliorating a medical condition -
XX
PS Disclosure: Page 51; 519pp; English.
XX
CC AAD05121-AA05203 represent cDNAs corresponding to 24 human secreted
CC protein genes, and AAE01232-AAE01311 represent the proteins they encode.
CC AAE01312-AAE01340 represent human secreted protein variants or fragments.
CC The secreted proteins and their genes are useful for preventing,
CC treating or ameliorating medical conditions, e.g., by protein or gene
CC therapy. Pathological conditions can be diagnosed by determining the
CC amount of the new protein in a sample or by determining the presence of
CC mutations in the new genes. Specific uses are described for each of the

```


CC 24 genes, based on the tissues in which they are most highly expressed,
 CC and include developing products for the diagnosis or treatment of
 CC proliferative disorders, cancer, tumours, foetal and developmental
 CC abnormalities, haematopoietic disorders, diseases of the immune system,
 CC AIDS, autoimmune diseases (e.g., rheumatoid arthritis), inflammation,
 CC allergies, neurological disorders (e.g., Alzheimer's disease,
 CC Parkinson's disease), cognitive disorders, schizophrenia, asthma,
 CC skin disorders (e.g., psoriasis), sepsis, diabetes, atherosclerosis,
 CC cardiovascular disorders, angiogenic disorders, kidney disorders,
 CC gastrointestinal disorders, pregnancy-related disorders, endocrine
 CC disorders, and infections. The proteins can also be used to aid wound
 CC healing and epithelial cell proliferation, to prevent skin aging due to
 CC sunburn, to maintain organs before transplantation, for supporting cell
 CC culture of primary tissues, to regenerate tissues, to identify their
 CC cognate ligands or binding partners, and in chemotaxis, and can be used
 CC as a food additive or preservative to modify storage properties.
 CC Antibodies specific for a protein of the invention can be used in
 CC alleviating symptoms associated with the disorders mentioned above, and
 CC in diagnostic immunoassays e.g., radioimmunoassay or enzyme linked
 CC immunosorbent assay (ELISA). The present sequence represents a human
 CC secreted protein fragment referred to in the disclosure of the invention.

XX Sequence 993 AA:

Query Match 3.2%; Score 235.5; DB 22; Length 993;

Best Local Similarity 18.8%; Pred. No. 2.8e-11;

Matches 218; Conservative 189; Mismatches 437; Indels 313; Gaps 52;

QY 97 CVAMNDKDDVLAV-IAEKSSCIYLMADANTNTSOLD-----NGMRDQ 139
 -DB 58 cva--yakdkrfagsadks--vllw-----tsklegllkythndaigcveynphtg 107
 QY 140 MS-----FLMKSQVSFLAVGTGKGLATYNNQTSRKIPVLGKHKRITCGGMAENLX 133
 -DB 108 laecssadslglws-----peqkvsxkhskskkl-----lccswtndgqyl 147
 QY 194 ALGCEPMKITVSNQEGDTIKOTQVRSEPMXNOFFLMKMDRTSAESMISVGLGKTLFF 253
 -DB 148 algmfmglislrnkgeekklterpgsalsplwsicmnspreerndilavdagqkvsfy 207
 QY 254 -LNLNPNPNADLEFOODFCGNIVCYNMGDGR-IMIGFCGHRFVISTHTGELGOETFOA 311
 -DB 208 qlsgkqlgkdrainfdp-----ccislyfkgeyillg--gsdkqyslfc----- 249
 QY 312 RNNKDNLTSAVSGTLNKVATCGDNCIKIODLVLDKDMYIILDEBNKGLGLSMTDGG 371
 -DB 250 --kdgv-----rlgtvg-----eqnswvwtcgakpds 274
 QY 372 QLLALSTQNGSLHVFLLKLPLIGDACSTRATVLTSLLEVTVANPVEGELPITVSVDPEPN 431
 -DB 275 nyvvvgcgqgclsfyqlifstvhnglykdyaydsmdvqlhlt-ekqvrkikckelyk 333
 QY 432 FVAVGYHLAVGNMNRAMPVVLGENAVKRLDMELYGLTVASICLHSDYAALFEKGVOHL 491
 -DB 334 klaiynrlaiqlpeklllyelys---edlsdmhy-----rvkekiklkfecnll- 381
 QY 492 LISEILDAQBER-ETRLPAVVDK-----CRILCHALTSPDLITGTGTGVQYF 540
 -DB 382 vcanhlllcekrllcqlsfisgvrerewmesllrylkivgpprgregllvgllknqgllk 441
 QY 541 YIEDMGFVNDYRHPVSVKRTFFPDPNGTRLVFEIDKSDGFFYCCPVNDAVTEIPPESPFIK 600
 -DB 442 fvdnlfaivllkqatavrcldmasarkklavdendtcly---didtckellfgepans 498
 QY 601 VLWENPMMDKGVFT--AYDDKVVYTVFHKDTQGAKEVILLAGSTKVPFAHKPLLLYNG-E 657
 -DB 499 vavnlqcedmlcfegggylilikastfpyhrqklg---fvvg-----yngsk 542
 QY 658 LFCOTQSGKVNNTYLSHGFLSNLKDGPDELPRMLAHNMLKRFSDAMEMCRILNDEAA 717
 -DB 543 lfc-----lhvf--sisaveypgsapmyy-ltdrkfkfkaaylaclgvtld 586

QY 718 WNELARACLHMEVEFAIRYRRIGN---VGIVMSLEQIK--GIEDYNLLAGHLAMFTND 772
 -DB 587 wrleamealeldefetackafirvgdlrylellssleerkkfgetnndilfaadvfygqk 646
 QY 773 YNLADLYLASSCPILAELMRDLQHWDSALQAKHLAPDOIPFI-----SKEVA 822
 -DB 647 fheaaklykrsgnenallemtyldcmfeyakdfllsgsdpckemllltkqadvarnlkepk 706
 QY 823 IQLEF---AGDYVNALAHYKGTGCKNKEDEKCLAGVADOMSTRMDDIRGVQALKHS 879
 -DB 707 aavemyisgehvkae-----lcgch-----gwdmll--diarkikdaerepl 749
 QY 880 RVLKRCGALTEEMKQFSEAAOLYEGKGLYKRAASYIRSKNMAK---VGDLPHVSSPK 936
 -DB 750 ll-----cayllkklidpyyaaeeyllmgdlksivqhlvetqrdeaalagkhpctkd 804
 QY 937 IHLQYKANEADGRYKEAVVAYENAKOMQSVIRIYLDHLNPNKAVNIYRETSGLGAKM 996
 -DB 805 lymypaqwlaendrfeeaqafhkgqrgeavqv-ldqtltn-----navaesrfndaa-- 856
 QY 997 VARFFQLGDYGAIOFLVMSKCNNAFTLAQOHNMK-ETIYADITSGEDTNTDYOSIAL 1055
 -DB 857 -----yywmstlscldiaqdpqcdmlgkty-----hfgrlae 891
 QY 1056 YFEGEKRYLQAGKFFLLCGYSRALKHFLLKCPSEEDNVAIEMAIETVGAOKDELFTN--Q 1113
 -DB 892 lylng-----yhaihrh-----tedpfswh-----rpetllfnlstr 920
 QY 1114 LIDHLLGENDGMPK-----AKYLFRLYMALKQYREAAGTAIIRAEOSAGNY---R 1163
 -DB 921 flhh-----slpkdcpsgjstkwkllftl-----akgskaigayrlar 957
 QY 1164 NAMDVLFMSWAEKLSOK 1180
 -DB 958 haydklrglylparfqk 974

RESULT 15

AA933308
 ID AAB93308 standard; Protein; 993 AA.

XX AAB93308;

XX 26-JUN-2001 (first entry)

XX Human protein sequence SEQ ID NO:12383.

XX Human; primer: detection; diagnosis; antisense therapy; gene therapy.

XX Homo sapiens.

XX EP1074617-A2.

XX 07-FEB-2001.

XX 28-JUL-2000; 2000EP-0116126.

XX 29-JUL-1999; 99JP-0248036.

XX 27-AUG-1999; 99JP-0300253.

XX 11-JAN-2000; 2000JP-0118776.

XX 02-MAY-2000; 2000JP-0183767.

XX 09-JUN-2000; 2000JP-0241899.

XX (HELI-) HELIX RES INST:

XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX MPI; 2001-318749/34.

XX Primer sets for synthesizing polynucleotides, particularly the 5602

XX full-length cDNAs defined in the specification, and for the detection

XX and/or diagnosis of the abnormality of the proteins encoded by the

THIS PAGE BLANK (USPTO)

QY 51 TWLAAPIQFAMQKTSNGNLAY-----TGADYIKIEDIRG--OKRSEINLPGNCVAMDM 102

199 TWRG-----DGGFAFVSYVCPETGAKK-VKRMREERALDOSTE-PVAGLGPLAW 246

Db 103 DKDDVLAVLIAEKSSCIYLMADPANTNKTSLDNGMRDMSEFLMWKGSFLAVGIVKGNLX 163

```

Db 247 KPSGSLIASTQDKP-----NOODIVFEKNG-----L 273
QY 163 IYNHOTSRKIPVLGKHTKRTTCGWNAE-NLXALGDEDMITVSNQEGDITROTQVRSER 221
Db 274 LHHFTF---LPFL-KDEYKNDLMMNADSSVLAVERED-----LORKESSIPYTCQVLM 324
QY 222 XNQPFELMKMDRTSAESMISVVLGKTLFELNLEPNPDLEFQODGNTVCYNWYG 281
Db 325 VGNVHMYLK-----QSLSFSTCGSKSLVSL-MMDPVLPYRLHLVLCQGMHVLAVDMWH 375
QY 282 DGRIMGFSGGHFVISTHGE-----LCQELFOARNKH 315
Db 376 TIDRSVGDNSDSLNVNAVIDGNRLVTVFRQTVVPPMCTYOLLFPHVNVQVFLAHPOK 435
QY 316 DNLTSIAVSQTLNKVAT--CGDNCKIQLDLY-----DLKDMVILNLD 356
Db 436 SN--DLAVLADASNOISYKGD-CPSPADPTVKGAVGSGFKVCLRTPHLEKRTKIQFEN 482
QY 357 EENKG-----LGTLSMTDQQLALS---TORGSLHVFLLKPLIGDACSTRIAVLTSL 407
Db 493 NEDQDVNPLKTLGLTMEEDVFLAVSHSPRSVTH-HLT-----AASSSEM----- 538
QY 408 LEVTVANPVGELPIYVSVDYEPNPAVGLYHLAVGNRNRAVTVLGENAVKRLKMEYL 467
Db 539 -----DEHGQLNVSSSAVD----- 554
QY 468 GTVASICLHS---DYAAALFEKVQLHLESEILDAOEERTRLP----- 510
Db 555 GVTISLCNKSRTSVVLQADGOIFKYLMSPSLAIKPKMNSGFFVRFPTQOTELAM 614
QY 511 -----AVDKCR-----ILCHALTSDFLIYGTDTGVVQYFIEDMQF--- 547
Db 615 IGEBCVLGTLTDCRFINDIEVASNITSFAYVDEFLLTTHSHTCQCFCLRDSFKTLQ 674
QY 548 -----VNDYRHPVSKKIFPNNGTRLYFIDEKSDGFVYCVNDAIYEIPDSFTIGVLM 603
Db 675 AGLSHHVSHGEVLKRV--ERGSRTVTV-VPODTKLYLQMPRGNELVNHRALVLAQI- 729
QY 604 ENMPMDKGVFIAYDDKVTVYVFNKDTIOGAKVILAGSTVPRNRPRLLYNGELTCOTQ 663
Db 730 RKM-LBKLMF-----KEAFECMRKRLINLNI-YDHNPKVFLGVEFFINQ 773
QY 664 SGKVNNTIYSTHGLSNLKDQXGDELKPM---LAHNLMLKPSDAMEKRIIINDEAMN 719
Db 774 IDSVNHINTL--FTELKE--EDVTKTMYPAVTVSSVYLSRDPDNKIDLVCDAARAVM 827
QY 720 ELA---RACL---HHM-----EVEFAI-RVYBRIGN-----VGIWMSLEDKIGIEDY 759
Db 828 ESTNPKYICLSITSHVKKTPPELETYVLQVHVLQSNAPSDDPDAVSAEELKTLHLVDY 887
QY 760 NLAGHILAMETNDYNIADLYLASSCPIALAEMRRDLQHMDSALQAKHLAPDQIFISK 819
Db 888 NELYDH-SLGTYDF---DLVL-----MVAEKQKDPKEVLPFLNLTKMETNYQFETID 937
QY 820 EYALQLEFAGDYVNAALHAEKGTGDNK--EHDEACLAGVAQMSIRMGDIRRGVQALX 876
Db 938 KY-----LKRREKAIGHLSKCGPEYFPECLNLIKDNL-----YNEALK 976
QY 877 --HPSRVLKRD---GALLENKQFSEAAQLEYKGIYDKASVYIRSKNNAKVDDL 930
Db 977 LKSPSSQOYODISIAVGEHLMQEHMTEPAGLIMARCAGHAHKAASLALTCGNW----- 1028
QY 931 HVSSPKIHQYAKAKKADGRYKEA--VVAYENAKQMSQVIRIYDLHLNNEKAVINIRET 988
Db 1029 -----KQALCVAAQNLNTKQVGL-----GRTLAGKLYEOR 1060
QY 989 QSLDGAKVYARFLLQDYGSAIQFLWMSKCNNEAFTLAQOHKMEI----- 1035
Db 1061 KHLIDAA-VYLEBCAQ--DYEEAVLLLEGAAMEALRLVYKYWRDLIETNVKPSILEAQ 1117
QY 1036 --YADIGSEDTT-----NEDYQSIALYFEGEKRYLDAGKFFL----- 1071

```

```

Db 1118 KNYMFLDSQATWATFSRHKRLVRLVRELKBOAQAGL--DDEVPHQGESDLFSETSSVSG 1175
QY 1072 --LCGOYSRA-----LKHFLKPSCEDNVATEMIETVGO-----AKDE 1108
Db 1176 SEMSGKYSHNSNRTSARSSKRNRRKARKHSLKEGSPLEDLALLESEVQNTENLKD 1235
QY 1109 LRTNOLIDHLGENDMPKDAKILFRLYMAKQYREAAQ---TATITIAI----- 1154
Db 1236 -----VHHL-----KVLV-LFEFDEQGRELQAFEDTLQLMERSLPEIWTLY 1278
QY 1155 EEOAGNRYNRADVLFSSMYELKSOKIKIP 1184
Db 1279 QONSATPVLGPNSTANSIMASYOQKTSVP 1308

RESULT 2
US-09-286-891-2
Sequence 2, Application US/09286891
Patent No. 6172195
GENERAL INFORMATION:
APPLICANT: Cohen, Lucy
APPLICANT: Baenlele, Patrick
TITLE OF INVENTION: IKAP Proteins, Nucleic Acids and Methods
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 75 DENISE DRIVE
CITY: HILLSBOROUGH
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94010
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/286,891
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/971,244
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: OSMAN, RICHARD A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: T97-011
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 343-4341
TELEFAX: (650) 343-4342
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1332 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-286-891-2

Query Match 2.3%; Score 167; DB 4; Length 1332;
Best Local Similarity 18.7%; Pred. No. 1.2e-05;
Matches 253; Conservative 186; Mismatches 455; Indels 456; Gaps 66;

```

```

Db 274 LHGHT---LPFL-KDEVKNVNDLMMNADSSVLAVRLED-----LOREKSSIPKTCVOLMT 324
QY 222 XNQOFLMKMDRRTSAESMISVVLGKKTLEFLNLNPNRDLERQODGNCVYCNWYG 281
Db 325 VGNVHMYLK-----QSLSESTCGSKSLVSL-MMDPVTYRLHVLCOGMHRLATIDMW 375
QY 282 DGRIMIGFCGHPVISTHTEGE-----LGQELFPOARNHK 315
Db 376 TTRSVYGDNSDLSNVAIDGNRVLTVPFQTVVPPMCTYQLLPHRPVQVFLNHPQK 435
QY 316 DNLSTAVSOTLKKVAT--CGDNCKITQDLV-----DIKMWVILNLD 356
Db 436 SN--DLAVLDASNQISVYKCGD--CPSSADPTVKGAVGSGFKVCLRPHLEKRYKIOFEN 492
QY 357 EENKG-----LGTSLMTDDGQLALS-----TORGSLHVFELKPLIGDACSTRILATVSL 407
Db 493 NEDODVNPRLGLLITMIEEVEFLAVSHSESPSPSVH-HLT-----AASSE----- 538
QY 408 LEVTVANPVEGELPITVSVYDEPNFVAVGLYHLAVGMNBNRPFVILGENAVKRLKDMEYL 467
Db 539 -----DEERQGLNVSSAAVD----- 554
QY 468 GTVASICLHS--DYAALEFGKVQHLHLESELDAQERETRLP----- 510
Db 555 GVIISLCNSKTSKVYLQDLADQOIFKYLWESPSPSLAIKPMKNSGCFVRFPYPCOTELAM 614
QY 511 -----AVDDKCR-----ILCHALSDFLITDTGCVQVFIEMQF----- 547
Db 615 IGESECVLGLTDRCKRFINDIEVASNTSPFVYDEFLLTTHSHTCQCFCLRASFTTLO 674
QY 548 -----VNDYRHPVSVKFIFFPNNGTRFLVFIDEKSDGEVYCVNDATYIEIPFSPITKGLM 603
Db 675 AGLSNHHVSHGEVLKRV--ERGSRIYTV-VPRDTKVLYDMPRGNLEEVNHRALVLAQI- 729
QY 604 ENRPMDKCFIAVDKDVYTVVYVHKDTQGAKVILAGSTKVPRANKRPLLNGELTCOTQ 663
Db 730 RKM-LDKLME-----KEAFECMRKRLINLNP-I-YDHNPKVGLVGVETFIKQ 773
QY 664 SGKVNNTVSTHGFSLNKDGPDELPRM-----LANHMLKRRPSDAMEMCRILIIDEAMN 719
Db 774 IDSVNHNINL-----FTELKE--BDVTKTMYPAVYTSVYLSRDPDGKIDVLCAMRAVM 827
QY 720 ELA---RACL---HHM-----EVERAI-RVYRRIGN-----VGIVMSLEOIKIEDY 759
Db 828 ESINPMKYCGLSILTSYHKKTPELEIYLOKVHLEQGNAPSDPDVSAEALKYLHLHVDY 887
QY 760 NLAGHLAMTNDYNYLAODLYLASSCPIALALEKRRDLOHMDSLDLAKHLAPQOIPISK 819
Db 888 NELYDH-SLCTYDF--DLVL-----MVAEKSOQDPKEYLPFLNLTLLKMETVYQRETTID 937
QY 820 EYAIQLEFAGDVYNALAHYKGTGNK--EHDEACLAGVAMQSTIMGDIRRGVNOALK 876
Db 938 KY-----LKRTEKALIGHLSKCGPEYFPPECLNLIKDKNL-----YNELAK 976
QY 877 --HPSRYLKRDG-----GALENMKOFSEAAQLYKGLYDYKASVYIRSKNMAKVGDLRP 930
Db 977 LYSPPSSQVQDIDIAVGENHLMQEHMEPAGLMEFARGCAHKKALSAFLTCGNM----- 1028
QY 931 HVSSPKIHLQYAKAKADGKRYKA--VVAENAKQWOSVIRITLIDLHNPEKAVNIYRET 988
Db 1029 -----KQALCYAADLNFTKDLVGL-----GRTLAGKIVEOR 1060
QY 989 OSLDGAKMVARFFLOLDGYSAIOPLVMSKCNNEAFTLAQOHNKMEI----- 1035
Db 1061 KHIDAA-MYLEBCAQ--DYEAVALLLBEGAMEBALRLVYKYNRDLIETNVKPSILEAQ 1117
QY 1036 --YADITGSEDJT-----NEDYOSIALYFESEKRYLQAGKFFL----- 1071
Db 1118 KNVMAFLDSOTATFYSRKKRLLVVRREIKBEQAQAGL--DDEVHGHQSDSEFSESSVYSG 1175
QY 1072 --LCGQYSRA-----LKHFLLCPSSSEDNVAIEMALETVQO-----AKDE 1108
Db 1176 SEMSGKYSHSNRISARSSKRNRRKAERKKHSLKEGSPLEDLALLEALSEVQNTENIKDE 1235

```

```

QY 1109 LNTNQLIDHLGENDGMPKDKYLFRLYMAIKQYREAQ-----TATIIAR----- 1154
Db 1236 -----VYHL-----KVLF-LFEPDEGRLOKAFEDTLOLMERSPEITWLY 1278
QY 1155 EEOAGNYRNADHVFESMVAELKSORIKIP 1184
Db 1279 QQNSATPVLGPNSTANSIMASYQQOKTSVP 1308

RESULT 3
US-08-775-009-35
; Sequence 35, Application US/08775009
; Patent No. 5935783
GENERAL INFORMATION:
; APPLICANT: Gong, Wellong
; APPLICANT: Emanuel, Beverly S.
; APPLICANT: Budarf, Marcia L.
; APPLICANT: Roe, Bruce
; TITLE OF INVENTION: No. 5935783el Genes Mapping in the Digeorge and
; TITLE OF INVENTION: Vellocardiofacial Syndrome Minimal Critical Region
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn, Kurtz MacKiewicz &
; ADDRESSEE: No. 5935783ris, LLP
; STREET: One Liberty Place - 46th floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: U.S.A.
; ZIP: 19103
COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/775,009
; FILING DATE: 27-DEC-1996
; CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
; NAME: Trujillo, Doreen Yanko
; REGISTRATION NUMBER: 35,719
; REFERENCE/DOCKET NUMBER: CH-0681
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3439
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 871 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-775-009-35

Query Match 1.9%; Score 140.5; DB 2; Length 871;
Best Local Similarity 19.0%; Pred. No. 0.0015;
Matches 158; Conservative 113; Mismatches 257; Indels 305; Gaps 39;
QY 445 NNRAMFVYLGENAVK-KLQMEYIGT--VASICLHSD--YAAALFE-GKVQHLIESEIL 498
Db 113 MALAKIYIDSNNNPRLRNPYDSVYGVKYEKRPDLACVAYENGQCDLELI----- 167
QY 499 DAOEBERTLFPVADKCRILCHALTSDFLIYGTDFGVQVYFIEMQFY---NDYRHP 554
Db 168 --NVNENSLSFKSL-----SRVLYRRKDP-----ELMGSVLLESNPYRRP 205
QY 555 VSVKKTFFPDNGRRLVFIDEKSDGFYVCPYNDATYELPDSPITIKVLMENMP-----M 608
Db 206 L-----IDQ-----VVQTAALSETQDPEEVSIVKAFMTADLNELETL 244
QY 609 DKGVFIAYDDKVVYTVVYVHNDITQGAKVILAGSTKYVFAHRPILLYNGELTCOTQSGKVN 668

```

Db 245 EKIIVL-----DNSVFSHRMILQMLITLTAIKADTRV-----MEXINRLD-NYDA-PDLA 292
QY 669 NIYLSTHGF-----LSNLKDXGPDELPRMLAHMLMKRPSDAMEKCRILINDEAAMELAR 723
Db 293 NAIISMELEFEBAIRKRDVNTSAVQVLEIHGNLDR---AYEFERENEPNVSOLAK 349
QY 724 ACLHMEVEFAIRVYRRIGNVIGVLSLEQIKGIEDYNLLAGHLAMFTNDYNLAQDLYLAS 783
Db 350 AOLQKGMVKEAI-----AOLQKGMVKEAI-----DSYIKA 367
QY 784 SCPIALAEHRDQ---HMSDA---LQLAHLAPDQIPFISKEYATQLEFAGDYVALAH 837
Db 368 DDPSSTMEYVQAANTSGNMEELVYKLOMARKKARES--YETE---LLEFALAKTNRLAE 421
QY 838 YEKGITGDNKEH-----DEACL-----AGVAQMSIRMGDIRGVNOAL 875
Db 422 LEEFINGPNNNAHQOYGDRCYDEKMDAKKLNNVSNFGRSLASTVHLGELYQAVDGAR 481
QY 876 K-HPSHYLKRDCGAILENK-----KANSFTWKEVCFACVDGKEFRLAQMGCHIVHADELELINYQDRGFEEELITMLBA 541
Db 482 KANSFTWKEVCFACVDGKEFRLAQMGCHIVHADELELINYQDRGFEEELITMLBA 541
QY 906 GLYVOKA-----ASYVIRSKN-----NAKVDDLPHV--SSPKIHL----- 939
Db 542 ALGLEBAHGMFTELAILYKFRPKQMRHELEFMSRVN--IPKYLRAAEQAHMAELVE 599
QY 940 QYAKAKKAD-----GRYKAVAVAYENAKOMOSVIRIY----- 971
Db 600 LVDKVEEYONAIITMMNHPTDAMKEGQFKDITIKVANVELYRALOYFDEKPLINDLL 659
QY 972 -----LDHLNPEKAVNIVRETQSIDGAKMYAR-----FLQIGDYOS 1009
Db 660 MYLSPLRLD---TRAVNFSKYKQPLVYKPYLRVQNNHNSVNSLNNLITTEEDY-Q 714
QY 1010 AIOFLVMSKCNNEAFLTAQOHNMKEIYADIIIGSEDTNEDYOSIALY-PEGEKRYOACK 1068
Db 715 AIRTSDAIDANDNISLARLEKHELI-----EFRIAIYELKGNKRMKQSAVE 762
QY 1069 FELLQGYVSRAHLKHLKCPSSSDNVAIEMAIFVGQAKDELTLIDHLIGE 1121
Db 763 ---LCKKDS-----LYKDAQVASESKTLELAELLQMFDE 796

RESULT 4
US-08-874-266-2
: Sequence 2, Application US/08874266
: Patent No. 5955279
: GENERAL INFORMATION:
: APPLICANT: Gatti, Richard A.
: TITLE OF INVENTION: ATAXIA-TELANGIECTASIA: MUTATIONS IN THE ATM GENE
: NUMBER OF SEQUENCES: 33
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Knobb, Martens, Olson and Bear
: STREET: 620 Newport Center Drive 16th floor
: CITY: Newport Beach
: STATE: CA
: COUNTRY: USA
: ZIP: 92660
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: OPERATING SYSTEM: DOS
: SOFTWARE: FASTSEQ for Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/874,266
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Wags Vensko, Nancy

REGISTRATION NUMBER: 36,298
REFERENCE/DOCKET NUMBER: DCIA006,006A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-235-8550
TELEFAX: 619-235-0176
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 3056 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
US-08-874-266-2

Query Match 1.9%; Score 136.5; DB 2; Length 3056;
Best Local Similarity 18.0%; Pred. No. 0.034;
Matches 236; Conservative 176; Mismatches 420; Indels 481; Gaps 62;

QY 323 VSOTLNKVAICGDNCKIKQ-----DLVLD-KDMYVL-----NLDEENKGLTSMTD 369
Db 455 VLKCLTEVALCODKSNLESQKSDLIKMKIKWCTTFRGISSEIOQAEISGL--LGALI 512
QY 370 DQOLLAISTQSLHVFYTKLPILIDACSTRYAYLTSLLEVTVANPVGELPITVSQVE 429
Db 513 QGSLVEVDREPKFLFTGSACRSPCAVCCLLALFTSI-----VGAVKMGIE 560
QY 430 PNFVAVGLYHLAVGNMNAFVLEGNMVKL-----KDMYLTGVASICHSYAAAL 483
Db 561 QNMCEV-----NRS--FSLKESIKWLEFYQLBQEDLENSTVEPPI-LHSNPHLY 607
QY 484 FEG-----KVQLHLIES--EILDAQERETRLPAVD-----DKCRILCHAL 523
Db 608 LEKILVSLMKKCKAMNFGQSPCEHHQKKELESSEVELLQTTFDK----- 659
QY 524 TSDFLYGTDTGVQY-----FYIE-----DMQFVDYRHPVSVKKIFDP 564
Db 660 -MDFLTIVRECIERKHOSIGFSVHQNKESIDRCILGLSEQLNNYSSEIT----- 710
QY 565 NCTRLVFLIDKSDGVYCCVNDATYEIDPFSPITGLVMEWMPMKGVFAVDDQKYTY 624
Db 711 -----NSETLVRC-----SRLVGLV-----GCT-----CYMG 733
QY 625 VFKHDTIOGAKYTLAGSTKVPFAHKKPLLYNGELTCQTSQKVNNIYLSHGFSLNLKDX 684
Db 734 VIAEEAVKSELFQKANSIMQCAGESITLFFKKTNDIEFRIGSLRMNMOCTCRLSNCRK 793
QY 685 GPDELRP-MLAHNMLKRESDAWEMCRILINDEAANMELARACLHMEVEFAIRYRIGN 743
Db 794 SPNKIASGFFRLRLSKLMDIADICKSL-----ASFIKKPPDR--- 832
QY 744 VGIWVSL-----QIKGIEDYNLLAGHLAMFTNDYNLAQDLYLASSCPALAEHRDQ 798
Db 833 -GEVSEMEDDTNGNLEVEDOS---SMNLF-NDY-----PDSSVSDANEQGE 874
QY 799 WDSALQAKHLAPDQIPFISKEYATQLEFAGDYVALAHYEKGTGDNKEHDEACL----- 854
Db 875 SGGSTIGAINPLAE-----YLSKQDLLFL-----DMLKFL-----CLCYTT 910
QY 855 AGVAQMSIRMGDIRG---VNOALKHPSR-----VLKRDGCA-----ILENM 893
Db 911 AQTNTVSFRAADIRKRLMLLIDSSYLEPTKSLHLMYLMLKELPEGEYPLPMEVDLELL 970
QY 894 KQFSEADQLEKGLYDKAASYIRSKNMKAKGDLPHV-----SSPKIHLQYAKAKAD 948
Db 971 KPLSNVCSLYR-----RDQVCKT--ILNHVLAVYKVLGGSSNDSENTRDAO 1015
QY 949 GRKRAVAVAYENAKOMOSVIRIYLDHLNPEKAVNIVRETQSIDGAKMYARFFLOLDGYG 1008
Db 1016 GQFLTVIGAF-----W-----HLTKERKYYISVNM-----ALVNCITLLEADYIS 1056


```

COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/204,764
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/883,534
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PR-0332 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 606 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: NEUTGNT01
CLONE: 1221143
US-09-204-764-3

```

Query Match	1.98	Score	136	DB	3	Length	606
Best Local Similarly	19.18	Pred	No.	0.002			
Matches	92	Conservative	77	Mismatches	200	Indels	112
						Gaps	19

```

QY 65 SGNLYAVGADIVYIFED-----RHQGRSEINLPQVCVAMDMDDKDDVLAIVA-----EKS 117
Db 69 SGEYIASGDSYCKLIMWDTQKHEHLKYGYOPFACKIKRIKIANTEBSKRIAVVGEGRKEG 128
QY 118 CIYMDANT-----NKSOLDNGMRDQ----- 139
Db 129 AVFLMDSGSSVGEITGNHKNVINSVDIKOSRPRLATGSDDNCAAFEGEPFKFPTVGDH 188
QY 140 ---MSFLMSKVSFLAVGTGVKGNLXIYVHOTSRKIPVIG---KHHRITCGOMNENLX 193
Db 189 SRFVNCVAFRSPGNRFPATASADGQIITYDGKIGEKYCALGSGKANDGGIYLAISMWDS--- 246
QY 194 ALGEDKMIIVSNQEGDTIROTQVRSEPKXMOFPLMKMDRTSA---AESMISVVLGKKT 250
Db 247 -----THLLSASGDDTSKIMDVSVNSVSTFPMGSTVLDLQOQGLCMQKDHLLSVSL-SGV 300
QY 251 LEFLNLNPEPDNADLEFODDFGNIVCYNNY---GDSRIMIGFSCGHFVYISHTTGELOE 307
Db 301 INYLDNRNPSKPLHY-IGKHSKISIOCLYHKKNGKSYITSGSHDGHINWDSGTGE--ND 357
QY 308 IFQARNHKDNLTSLAVSOTLNRKATCGDNCKIRIODLVLDKM--YVILNLDEENKGL--- 362
Db 358 SFAGSGHTNOVSRLMTVDESGOLCISCMDDTVRYTSLM-LRDYSGQGVVLLQVQPKCVAVG 416
QY 363 ---GTLSTWIDGGLLSTQROGSLHVFLLTFLPLIGDACSRIATVLSLLEVTANPVEGEL 420
Db 417 PGGAIVAVVCIGIOTVLLKDR-----KC-----FSDNP--GTE 447
QY 421 PITSVDEPNEPVAAGLYHLAVGMNNRAWFY-VLGENAVAKKLDMETLGTVASISLHSDY 479
Db 448 PEYVAVHHGGRPVALG-----GVDSGNRLYSILGTLLDEGKLELAKGPVTDVAASHDG 501
QY 480 A 480
Db 502 A 502

```

RESULT 7
US-08-287-959-1

```

: Sequence 1 Application US/08287959
: Patient No. 5639651
:
: GENERAL INFORMATION:
: APPLICANT: Weissbach, Lawrence
: APPLICANT: Bernards, Andre
: APPLICANT: Settleman, Jeffrey
: TITLE OF INVENTION: GAP-RELATED GENE
: NUMBER OF SEQUENCES: 26
:
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Fish & Richardson
: STREET: 225 Franklin Street
: CITY: Boston
: STATE: MA
: COUNTRY: U.S.A.
: ZIP: 02110
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
:
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/287,959
: FILING DATE: August 9, 1994
: CLASSIFICATION: 514
: ATTORNEY/AGENT INFORMATION:
: NAME: Clark, Paul C.
: REGISTRATION NUMBER: 30,162
: REFERENCE/DOCKET NUMBER: 00786/181001
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (617) 542-5070
: TELEFAX: (617) 542-8906
: TELEX: 200154
:
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1657 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
:
: MOLECULE TYPE: protein
:
: PS-08-287-959-1

```

Query Match	1.88;	Score	133;	DB 1;	Length	1657;
Best Local Similarity	21.28;	Pred. No.	0.024;			
Matches 123;	Conservative	92;	Mismatches	244;	Indels	120;
					Gaps	28;

```
QY      663   QSGVNNNYLSTHGFSLNKKDCHGDEPLRPLAHNLMLKRFEDAMEMORILLUNDEANMELA    722
       : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db     468   ESDGVNTVMKOLSSSVGLTINEENCORYLEDTLMKLKAQAHAEN-----NEFTWMDI-    521
       : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY      723   RACHHHEVEFEALRYVRIGNVGTMSLEQTIGIEDYNLAGHLAMEFNDNLADDTLYLA    782
       : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db     522   QACYVDH-----VNLVGOEHER-----ILAGLINEALDEGAQKTLOA    560
       : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY      783   SSCPIALE--MRDDLOHWSALOAKHLAPDOIPTLSKEVAIOLEFRAGDYVNAHYATEK    840
       : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db     561   LQIPAAKLEGVLAEVAOHYODTLIRARR-----EKAOIDESEA---VIMLDEIQG    608
       : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY      841   GITDNKNRHEDA--CLAGVAQM--SIRMGDIRRGVNOALKPRSYLK---RDCAILENM    893
       : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db     609   GIMOSNKDOTOEAQCFALGIFAINAEVSGDVCKTLLS-ALRPDVGLTGYVPECCETH--    665
       : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY      894   KQSEAAQLYEKGUYLDRAASVYIRSKNMMAKVGDLLPHVSSPKTHLOYAKAREADGRYKE    953
       : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db     666   ---SDLBAKKKTLAVGDNNNSKWY--KHMYWG-----YYYYHNLETQEGWDE    709
       : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY      954   AVVAAYENAKOMQSIVIRIYLHDHLNPBEKAVNIY--REFQSLDGAMVAF-----FIOLG    1005
       : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db     710   PRNFVQNSMQ-----LSREIQSSISGVTAAVNRQOJMLNEGILTRLOARCRGVLVRO    763
       : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY      1006   DYGSAIQFLVSKSCNNNAFTLAQQ-----HNKMELIADIISSEDITPND--YQSIATLFE    1058
       : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db     764   EFERSRMFL--KOQIPATICTOSOMRGYKOKKAYVODLAYLRSHKDEVKVQTSIARNHQ    820
```

820

```

0Y 1059 GKKRYLQAGKFFLLCGYSALNHLFLKCPSSSEBNVAIEMAITVGOARD-ELTTHQMLDH 1117
      ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 821 ARKRYEYRLQYFRDHINDIIRKIAQIFRANKARD-----YTLTINAEDEPPVAVVRKVVH 874
      ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
0Y 1118 LIGENDGMPDAKFLFELYALVKQYBEAQAOTAIIRAESQAGNRNADHLYFSMYAEK 1177
      || : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 875 LLDQSD---DQ---FOEELDLKKMKREEVYITL---RSQQLQENDLNDIMKIRGL--Lv 921
      || : : : : : : : : : : : : : : : : : : : : : : : : : :
0Y 1178 SOKIKIPSEMATNLMLHSYILKIVKKNGDHKKGARM. 1216
      || : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 922 KKKTTTLD-----VYSHSKKLLK---KMKQQLSDMMMI 951
      || : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RESULT      8
US-08-508-836A-8
; Sequence 8, Application US/08508836A
; Patent No. 5777093
; GENERAL INFORMATION:
; APPLICANT: Shiloh, Yosef
; APPLICANT: Taglie, Danilo A.
; APPLICANT: Collins, Francis S.
; TITLE OF INVENTION: Ataxia-Telangiectasia Gene
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Reising, Ethington, Barnard & Perry
; STREET: P.O. Box 4390
; CITY: Troy
; STATE: Michigan
; COUNTRY: US
; ZIP: 48099
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/508,836A
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Kohn, Kenneth I.
; REGISTRATION NUMBER: 30,955
; REFERENCE/DOCKET NUMBER: P-313 (TAU)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (810) 689-3500
; TELEFAX: (810) 689-4071
; INFORMATION FOR SEQ ID NO.: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3056 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-508-836A-8

Query Match          1.8%; Score 131.5; DB 1; Length 3056;
Best Local Similarity 17.6%; Pred. No. 0.098;
Matches 231; Conservative 178; Mismatches 425; Indels 477; Gaps 60.

Oy   333 VSQTLNKVATGCDNCIKIQ-----DLVDLKDMYVLNT-----DEENKGLTSLSWTDG 371
      | : ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db   455 VLRCLEVALACODKRNSLESOKSDLLKMKIMWCTTFRGISSEQKQANFGILGAIIIG 514
      | : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Oy   372 QLLALSTGRGSIAHVELTKPLPIIGDACSRIATLTLSLEYTVANPVEGELPIYVSUVVEN 431
      | : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db   515 SLLEVVDREHWKLFETGSACRPPSCPAVCCLTALTTSI-----VPGVKKGIIEQN 562
      | : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Oy   432 FYAVAGLYHLAVGMNRRAMEYVLGENAVKKL-----KIDMEYLGTVASICLHSDYAAALPE 485
      | : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db   563 MCEV-----NRS--FSLKESITMKWLRLVQLLEGDLNSTEVPRI-LHSNFPHLVLE 609
      | : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Oy   486 G-----KVQHLHIES--ELIDAOEERTLRLEPAVD-----DKCRILCHALTS 525
      | : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

Dd	610	KILVSLIMKCKAAMNPFQSVRSEHHNKDBELSEVLEFLQDTTDDK	-----M	660		
Qy	526	DELIYGTDTGVVOY	-----FYIE-----	DMQVNDYRHPVSKYKIPDPNG	566	
Dd	661	DLFLTVRECGIEKHOSSIGFSVHQNKLKESJDLRCLLGSBOLLNVSSELT	-----	710		
Qy	567	TRLVFIDEKSGRVCYCVNDATYEIYDFSTIKGVLMEMWPRDKGVFLAYDDKAYTYTVF	-----	626		
Dd	711	-----NSETLYVRC-----	SKLIVGLV-----	CCY-----	CUMGVI	735
Qy	627	HKDTIOGAKVYLGASTVPAWPAHPRLLYUNELCYQOSQKNNVYILSTHGFSLNKDXGP	-----	686		
Dd	736	AEEBAVYSELPORANSIMQACAGESITLFEKKTNEEFRISSLKNNMOLTRCLSNCTKSP	-----	795		
Qy	687	DELBP-MLAHMLMKRPSDAMEMCRILUNDEAAMNELLARACLHMEVEFARIYVRRIGNVG	-----	745		
Dd	796	NKIASGFRLTLTKSLMNDIADICKSL	-----	ASFIKPFDR-----	G	833
Qy	746	IYMSLE-----QIKGIEDYMLLAGHLMFTNDYNLAODLYLASSCPIALALEKRRLOLWD	-----	800		
Dd	834	EVESEMEDTNGNLMVEVDOS-----SMNLF-NDY-----	PDSSVSPANPESGQ	876		
Qy	801	SALOLAHKAPDOLPRTFSKRYAQLQLEFAGOVVALAHYEKGTGONKEHDECL	-----AG	856		
Dd	877	STIGAINPLEE-----YLSKODLLFL-----	DMKFL-----	CLCVTTAQ	912	
Qy	857	VAQMSIRMGDIRRG-----VMAOLKHPSR-----	VLRDQCA-----	ILENNKO	895	
Dd	913	TNTVSPFRADIRRLMLLIDSSSTLEPTKSLHLMYLMILKELRGEBEYPLPMEDVELLKP	-----	972		
Qy	896	FSEAAOLEKGLYDKAASVYIRSKMAKVGDLPHV-----	SSPKIHLQYAKAEADGR	950		
Dd	973	LSNVCSLYR-----	BDODVCKT-----	ILNHVLHVYKNLQCSNMDSENTRDAGQ	1017	
Qy	951	YKEAVVYENAKQKQSVIRIYIDHLLNPRKAVNIYETGOSIDAKMVAFFLOLGDYSA	-----	1010		
Dd	1018	FLTVYIGAF-----W-----	HLTKERYTISVM-----	ALVNCILKTLLEADPYSKW	1058	
Qy	1011	IQFIVMSK--CENNAFT--LAOONHMKETIADIIGSEDTTNEDOSIALYFEGEK	-----R	1062		
Dd	1059	ALIVWKGDPFVNVFLFOPLADNHQVRYMLA-----	ASEINLFLPDRFGDSSR	1106		
Qy	1063	YLQA-----GKFLLQGOYSRALKHFLKCPSEDN-----	VAIMEALET	1101		
Dd	1107	LKALPRLKLOOTAFENAYILKAOGBRMSHSAENPETLDEIYNRKSVLLTLLIAVVLSCSP	-----	1166		
Qy	1102	VGO-----AKDELLNLOLIDHLGG-----	NGMPRDKAYLERLVMALK-	1140		
Dd	1167	ICEQOALFALCKSVYKENGLEPHLVKVKLEYVSETFSYRRLEPFMASHLDLYLEVLNLOD	-----	1226		
Qy	1141	-QYREAAOTATIIIREBQSGAGNYRNAHD-----	VLFSSYABLSQKIKPISE	1168		
Dd	1227	TEYNLSSPFLIL-----NTNIEDPFRSCYKVLPHLVIRSHFEDVKSIANOIQED	-----	1278		
Qy	1187	MATNMLMLSHYILVKI-----	HVXNGDHMGARML-----	1216		
Dd	1279	WKSLLTDCFPRIILVNLIPRAYEGTRDSGMAQORETAFKUYDMLKSENLLGQIDHFLIS	-----	1338		
Qy	1217	-----	IRAAUNISKFPHSIYIILTSTYIECHR	1243		
Dd	1339	NLPELIVELLMTLHPANSSASOTDLCDFSGDGLDAPRPNPSPHVYKATPAVYSNCHK	-----	1368		
Qy	1244	AGLKN-----SAFSFAMLMR-PEYRSKIDAKYKK-----	KIE-----	1275		
Dd	1399	TKLSILEILSKSPDSQIKILLALCEBOAETNNVYKKNHKLKITYHLFVSLLLKDLKSLG	-----	1456		
Qy	1276	---GWRPRDISIEEATTPCPC-----	KEFLPECELLCPGCKNSIPYCIATGRML	1325		
Dd	1459	GAMAFVLBDVYITLIHYNQRPSCIMOVLSRPSLC-CDLLSQVOCQATVYC-----	-----	1509		
Qy	1336	KDDMTVCPCHDFPVLVYSELKIMLMTBSTCPCMSERLMAAOLKKSIDCTOYL	-----	1376		
Dd	1510	KD-----ALENHLHVIYGT--LPLVYVEQOEVQ--QOVIDLLKYL	-----	1545		

; TITLE OF INVENTION: SCREEN FOR A PARTIAL A-T PHENOTYPE
 ; FILE REFERENCE: 22900033
 ; CURRENT APPLICATION NUMBER: US/08/642,274D
 ; CURRENT FILING DATE: 1996-05-03
 ; NUMBER OF SEQ ID NOS: 220
 ; SOFTWARE: Patent In Ver. 2.1
 ; SEQ ID NO 3
 ; LENGTH: 3056
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-08-642-274D-3

Query Match 1.8%; Score 131.5; DB 4; Length 3056;
 Best Local Similarity 17.6%; Pred No. 0.098;
 Matches 221; Conservative 18; Mismatches 425; Indels 477; Gaps 60;

QY 323 VSQTLNKVATCGDNCIKIQ-----DLVDLKMVYIINL-----DEBNKGLGTLMSWTDG 371
 DB 455 VLRLCTREVALCQDKRSNLESSQKSDLKLNKIMCITFRGISSEQKQAEKFGLLGAIIOG 514
 QY 372 QLLALSTQKSLHVFELTKLPILDACSTRAYLTSLELVANPVEGELPITYSVDEPN 431
 DB 515 SLVEVDREFWKLFTGSACRSPAVCCLTALTSTI-----VGAVMGKIEON 562
 QY 432 FVAVGLYHLAVGNNAFVIGENAVKRL-----KDMELGTVASICLHSDYAAALFE 485
 DB 563 MCEV-----NRS--FSLKESTIKKLLFQLEGDLNENSTEVPTI--LHSEPHLYLE 609
 QY 486 G-----KVQHLIES--EILDAEERETRLFPAYD-----DKCRILCHALTS 525
 DB 610 KILVSLTMKCKAAMNPFQCEPHHKKDELSFSEVELFLQTFEDK-----M 660
 QY 526 DELIYGTDRGVQY-----FYIE-----DMQFVNDYRRPVSVKIIFDPNG 566
 DB 661 DFLTVIRECEIEKHQSSIGSVHQNLESIDRCLLGISEQLNNYSEIT----- 710
 QY 567 TRLVFIDEKSDGFVYCPVNDATYEIPDFSPITKIVLMMNPMQGYEIVADDKVTYV 626
 DB 711 -----NSETLVRC-----SKLLGVV-----GCY-----CIVGVI 735
 QY 627 HKDTQAKAYIAGSTKVPFAHRLPYLNGELTCQTSQSKVNNIYLSHGLSNLMDXG 666
 DB 736 AEEBAVKSSELFQKANSIMQAGESITLFKNKTNEFFRIGSLRMNMOCTGCTSLNCTKSP 795
 QY 687 DELRP--MLANNMLKRFSDMWMCRILNDEAANNELARACLIHMEVEFAIRVRRIGNVG 745
 DB 796 NKIASGFFLRLLTSKLMNDIADICKSL-----ASFIRKPPDR---G 833
 QY 746 IVNSLE-----QIKGIEDVYLLAGHLAMFTNDYLAQDLYLASSCPILAEMRRDQHMD 800
 DB 834 EVSMDQDTGNLMEVEDOS-----SMNLF--NDY-----PDSSVSDANEPESQ 876
 QY 801 SALQALKHAPDOIPIFSKEKAYIQLEFAGDYVNAHAHYEGITGDKNEDEACL---AG 856
 DB 877 STGAINPLAEE--YLSKODLLFL--DMLKFL-----CLCVTTAQ 912
 QY 857 VAQMSIRMGIRRG-----VQAALKHPSR-----VLRDQGA-----ILEMKQ 895
 DB 913 TMTVSFRRAADIRKRLMLIDSSITLPEPKSLHLMYLMLEKELGEEVPLRMEVDYLELKP 972
 QY 896 FSEAAQLYEKGLYDKAASYIRSKNMAKVGDLLPHV-----SSPKIHLQYAKAKEDGR 950
 DB 973 LSNVCSLIR-----RDQDYCKT--ILNHVLAHYVKNLGSNMSMSENTRDQOG 1017
 QY 951 YKEAVVAAYENAKQWQSVIRIYLDHLNPEKAVNIIVRETQSLGAKVYAREFLQDGYSA 1010
 DB 1018 FLTVIGAF-----W-----HLTKERKRYFSVRM-----ALVNLCKTLLEADPYSKV 1058
 QY 1011 IOFLVMSK--CNNEAFT--LAQOHKKMEIYADITIGSEDITNEDYQSLALFEEGK-----R 1062
 DB 1059 AILNVMGKDFPVEVEVFOFLADNHQVRMLA-----AESINRLFLQDTKGDSR 1106

QY 1063 YLQA-----GKFFLLGQYSRALKHFLKCPSESDN-----VAIEMALET 1101
 DB 1107 LKALPLKIQDTAFENAYLKAQCGMRMESHSAENPETLDEITYRKRSVLLTILVAVLSCSP 1166
 QY 1102 VGO-----AKDELLTNQILIDHLGE-----NDGMPKDAKYFLRYALAK- 1140
 DB 1167 ICEKQALFALCKSVKENGLEPHLVKKVLEKVSFTFGYRLDEDMASHLDVYLEMNLQD 1226
 QY 1141 -QYREAQRAIITAREQSGNRYNAHD-----VLFSAYELKSKIKIIPSE 1186
 DB 1227 TEYNLSFPFILL-----NYTNIEDFYRSCYKVLIPHLVIRSHFDEKYSIANQIQED 1278
 QY 1187 MATNMLHSHYILVKI-----HVKNGDHMKGARML----- 1216
 DB 1279 WKSLLDLDCFPKILVNLIPRYATYGTSDSGMAQORETATKYMDLKENLGRQIDHLFIS 1338
 QY 1217 -----IRVANNISKEPSSHIVPILTSVTIECHR 1243
 DB 1339 NLPEIVELLMTLHEPANSASQSTDLCDPSGDLDPAPNPHPSPHVIKATFAYISNCHK 1398
 QY 1244 AGAKN-----SAFSPAALMR--PEYRSKIDAKYK--KIE----- 1275
 DB 1399 TKLSLLEILSKSPDSYOKTLAICEQAAETNNVYKRRILKIYHLFVSLLDKISGLG 1458
 QY 1276 ---GMYRRPDISEIEEATFPCCPC-----KFLPECELLCGCKNSIPCIATGRHML 1325
 DB 1459 GAAAFULRDVITYTLINHYNRPCINDVLSRSLC--CDLLSVOCTATAYTC----- 1509
 QY 1326 KDDWTVCPHCDPFPALYSELKIMLNTSTPCWCSERLNAOQLKISDCTOYL 1376
 DB 1510 KD-----ALENHLHIVGT--LIPLYVEOVEVQ--KOVLDLKYL 1545

RESULT 11

; US-08-952-127-3
 ; Sequence 3, Application US/08952127
 ; Patent No. 6211336
 ; GENERAL INFORMATION:
 ; APPLICANT: Shiloh, Josef
 ; APPLICANT: Tagle, Danilo A.
 ; APPLICANT: Collins, Francis S.
 ; TITLE OF INVENTION: ATAXIA-TELANGIECTASIA GENE
 ; NUMBER OF SEQUENCES: 24
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Kohn & Associates
 ; STREET: 30500 No. 6211336thwestern Hwy., Suite 410
 ; CITY: Farmington Hills
 ; STATE: Michigan
 ; COUNTRY: U. S.
 ; ZIP: 48334
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/952,127
 ; FILING DATE:
 ; CLASSIFICATION: 530
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Kohn, Kenneth I.
 ; REGISTRATION NUMBER: 30,995
 ; REFERENCE/DOCKET NUMBER: 2290,00029
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 810-539-5050
 ; TELEFAX: 810-539-5055
 ; INFORMATION FOR SEQ ID NO: 3:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 3056 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein

```

OY      1141 -GYRAAGCAIIIAAREEGSAGVNRNAH-----VFESMYAEKSOIKIPSE 1186
           ||| | :|| | :|| | :|| | :|| | :|| | :|| | :|| | :|| |
Db      1227 TEYNLSSTPFILL-----NYTNIEDPYRSCRYKLIVLVIIRSHFDEVKSIANDIOED 1278
           ||| | :|| | :|| | :|| | :|| | :|| | :|| | :|| |
OY      1187 MATNMLIHSYLVIKI-----HYKNGDHMKGARUL----- 1216
           ||| | :|| | :|| | :|| | :|| | :|| | :|| |
Db      1279 WKSLLTDFPKILLVNILPEFYAETGRDGSMAOOERATAYKVIMLKSENLIQKDHLFLFS 1338
           ||| | :|| | :|| | :|| | :|| | :|| | :|| |
OY      1217 -----IRANNISKFPPIHVILNSTVAIECHR 1243
Db      1339 NLPETVELLMTLHPDPANSSASOSTDLCDFGSDLDOPAPNPHPSPHVIKATPAVISNCBK 1398
OY      1244 AGLKN-----SAESPAMLMR-PEYRSKIADKYKK---XIE----- 1275
           ||| | :|| | :|| | :|| | :|| | :|| | :|| |
Db      1399 TKLKSILEILTSPPSYOKILLALICEQAEMTNVYKKRIKLITVLHFVSLLKDIKSGIG 1458
OY      1276 ---GWRRPDISIEBATTCPFC-----KFLLPCECLCBGCKNSIPYCIAIGRHML 1325
Db      1459 GAMAFVLPRVITYTLTHYINQRPSCTMDYSLRGFSLC-CDLLISOVCTAAVTYC----- 1509
OY      1326 KODWVCPCPHCDPALYSELKIMLNTESTPCMCSERINAOLAKKISPOTYL 1376
           ||| | :|| | :|| | :|| | :|| | :|| | :|| |
Db      1510 KD-----ALENLHLVIGT--LIPLVEOVEVO--KOVDLCLKYL 1545

RESULT 12
US-08-952-014C-3
Sequence 3, Application US/08952014C
Patent No. 6265158

GENERAL INFORMATION:
Applicant: Shiloh, Yosef
Title Of Invention: ATAXIA-TELANGIECTASIA GENE AND ITS
TITLE OF INVENTION: GENOMIC ORGANIZATION
NUMBER OF SEQUENCES: 91
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kohn & Associates
Street: 30500 No. 6265158thwestern Hwy., Suite 410
City: Farmington Hills
State: Michigan
Country: U.S.
ZIP: 48334

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
Computer: IBM PC compatible
OPERATING SYSTEM: PC-DOS/Ms-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/952.014C
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Kohn, Kenneth I.
REGISTRATION NUMBER: 30,995
REFERENCE/DOCKET NUMBER: 2290,00028
TELECOMMUNICATION INFORMATION:
TELEPHONE: 810-539-5055
TELEFAX: 810-539-5055
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 3056 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
US-08-952-014C-3

Query Match          1.8%; Score 131.5; DB 4; Length 3056;
Best Local Similarity 17.6%; Pred. No. 0.098;
Matches   231; Conservative 178; Mismatches 425; Indels 477; Gaps    60;
```

```

OY 323 VSQTLNKVATCGDNCKIQ-----DLVDLKDMYIINL-----DEENKGLGTLSTWDDG 371
DB 455 IURCLTEVALCQDKRSNLSOSSKSDLKLMNKIWCITTFPGISSEQQAENFGLLGAIIG 514
OY 372 QLLASTQSGSLHVLTKLPILIGDACSTRIAVYLSLEYVANPVGELPTIVSVDEPN 431
DB 515 SLVEVDREFWKLTFTGACRPSCPAVCLTLALTSTI-----VPGAKMGIEON 562
OY 432 FAVAVGLHVLAVGNMNAFVYGENAVKTL-----KDMELYOTVASTICHSQYAAALFE 485
DB 563 MCEV-----NRS--FSLKESITMKWLFYQLEGDLENSTVEPPT--LHSMFPHLYE 609
OY 486 G-----KVQLHLIES--EILDQERETRLPAVD-----DKCRILCHALTS 525
DB 610 KILVSLTMKNCAAMNFPSVPECEHHHNDKEBLSFEVBELFQTFPK-----M 660
OY 526 DELIYGTDTGVOY-----FYIE-----DMQVNDYHHPVSVKIPDPNG 566
DB 661 DELIYREGGIEKHOSIFSVHQNLIKESIDRCLLGLSEOLLNNYSSEIT----- 710
OY 567 TRLVFIDEKSDGFVPCVPVDATEIPEIDFSPITKGVLMEMPMKGVFIAYDDOKVITYE 626
DB 711 -----NSETLVRC-----SRLLGVV-----GCY-----CYGVI 735
OY 627 HKDTIGAKVILLAGSTKVPFAHKPLLLYNGELTCQTSQGVNNIYSTHGLSLNLDXGP 686
DB 736 ABEBAKSELPQKANSIMOCAGSITLFKNKTMEFRIGSLRMMQLCRLCSNCKRPS 795
OY 687 DELRP--MLAHNMLKRFSDAMEKCRILNDEAAMNELARACLHMEVEFAIRYVRIGNVG 745
DB 796 NKIASGFPLRLTSLKIMNDIADICKSL-----ASFIKKPPDR-----G 833
OY 746 IYWSLE-----QIKGIEDPILNLAHLAMFTNDVNLADLYLASSCPILALEMRDQHM 800
DB 834 EYESMEDDTNGNMEVEDS-----SMALF--NDY-----PDSSVDANEGESQ 876
OY 801 SALQIAKHLAPDOIPIFSKEYAIOLEFAGDYVALAHAYEKITGDKNHDEACL-----AG 856
DB 877 STIGAINPLAEE-----YLSQODLFL-----DMKFL-----CLCTYTAQ 912
OY 857 VAOASTIRMGDIRG-----VNOALKHPSR-----VLKRCGA-----ILENMO 895
DB 913 TMTVSFRADIRRLKMLIDSSSTLEPTKSLHLMYLMLEKELPEGEYPLPMEVDLELKP 972
OY 896 FSEBAOLYKGLYDKAASYIRSKNMKAGDILLPHY-----SSPKIHLOYAKAKRADR 950
DB 973 LSNVCSLYR-----RDQVCKT--ILNHVLHVKNLGSNMDSMTRODQO 1017
OY 951 YKEAVVAVENAKOMOSVIRIYDLHNNPEKAVNIVRETOSLDGAKVAREFLQDGYSA 1010
DB 1018 FLTVIGAF-----W-----HLTKERKIYSVRM-----ALVNCIKTLLEADPSKW 1058
OY 1011 IOFLVMSK--CNNEAFT--LAQOHKMEIYADITIGSEDTTNEYOSTIALYFESKEK----R 1062
DB 1059 AILNVAGKDFPVNEVFTOFLADNHQVRLMA-----AESINRLFOOTKGDSSR 1106
OY 1063 YLOA-----GKFELLCGYSRALKRFLKCPSEDN-----VAIEAIEF 1101
DB 1107 LLAALPLKLOQTAFENAYVLAQSGMREMSHSAENPETLDEITYRKSVLLTLVAVJSCSP 1166
OY 1102 VGO-----AKDELLTNQLIDLHLGE-----NDGMPKDAKYLFRLYMALK 1140
DB 1167 ICEKQALPALCKSVKENGLEPHLVKYLEKVSFTFGYRLREDMASHLDVLEMLNLOD 1226
OY 1141 -QYREAAQTAITAREBQSGNVRNAHD-----VLFPMYAEIKSOKIKIPSE 1186
DB 1227 TEYNLSSEFPILL-----NYTNIEDFYRSCYKVLPHLVIRSHDEVKSIANOIOED 1278
OY 1187 MATNMLILHSYILVKI-----HKVKGDMKGRML----- 1216
DB 1279 WKSLTLDCKPFLVNLIPFYAYEGSTRDSGMAQORETATKYMDIKSENLGKOIDHLFIS 1338
OY 1217 -----IRVANNISKFPSPHIVPILTSTVIECHR 1243

```

```

DB 1339 NLPETVELLMTLHEBPANSASQSTDLCDPSCGLDBAPNPHRPPIVATFAYISNCKR 1398
OY 1244 AGIKN-----SASFAMLMR--PEYRSKIDAKYK-----KIE----- 1275
DB 1399 TKLSITLETLSKSPDSYQYQILAIICEQAAETNNVYKRRILKIYHULFVSLLDIKSIGL 1458
OY 1276 ---GWRPRPDISEIEBATTPCPC-----KFLPECELICGCKNSIPLYCIATGRHML 1325
DB 1459 GAMAFVLRDVIYITLHYINORPSCIMDVLSRFSILC--CDILSQVCTQATYTC----- 1509
OY 1326 KQDMVTCPHQDPFALYSELKIMLNTSTPCPMCSERLNAOLKRIISDCTOYL 1376
DB 1510 KD-----ALENHLHVIGT--LIPLYEYQVEYQ--KQVLDLKLKYL 1545

RESULT 13
US-08-190-802A-66
; Sequence 66, Application US/08190802A
; Patent No. 5519003
; GENERAL INFORMATION:
; APPLICANT: Mochly-Rosen, Daria
; APPLICANT: Ron, Dorit
; TITLE OF INVENTION: WD-40 - Derived Peptides and Uses
; TITLE OF INVENTION: Thereof
; NUMBER OF SEQUENCES: 265
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Denlinger & Associates
; STREET: P.O. Box 60850
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306-0850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/190, 802A
; FILING DATE: 01-FEB-1994
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Fabian, Gary R.
; REGISTRATION NUMBER: 33,875
; REFERENCE/DOCKET NUMBER: 8600-0139
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ. ID NO. 66:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 514 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: YCW2 PROTEIN, Fig. 49
; US-08-190-802A-66

Query Match 18%; Score 131; DB 1; Length 514;
Best Local Similarity 20.9%; Pred. No. 0.0043;
Matches 89; Conservative 57; Mismatches 164; Indels 116; Gaps 20;

OY 4 IGRCSGLLDGNAVACIVETGACAVLHSSRGSSVEMKRIFSLKXTWLGAPIORAMOK 63
DB 132 VTRSSSAIGHGSTIIC-----SAFAPHTSSR----- 158
OY 64 TSGNYLAIVGA--DIYKIFDRHQKRSNLELPGN-----CVAMDQKDGVLAVIAEKSS 117
DB 159 -----MYTAGDNTARIMDCDTOTPMH--TLKGHYMWVLCVS--WSPDGEVIAT--GSMON 208

```


SEQUENCE CHARACTERISTICS:
LENGTH: 514 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: YCW2 PROTEIN, Fig. 49
US-08-473-089-66

Query Match 1.8%; Score 131; DB 4; Length 514;
Best Local Similarity 20.9%; Pred. No. 0.0043;
Matches 89; Conservative 57; Mismatches 164; Indels 116; Gaps 20;

QY 4 IGRCSGGLGDNVAVCLVETGACVCHSSRGSSVEMKRIFSLKRTWIGAPIQFAMOK 63
DB 132 VTRSSAIGHGSTILC-----SAFAPHTSSR----- 158
QY 64 TSGNYLAVTGA-DYIVKIFDRHGOKRSEINLPEN-----CVAMDMDKGDVLAVIAEKSS 117
DB 159 -----WVTGAGDNTARIMDCDTOTPMH-TLKGYMMVLCVS--WSPDGEVIAT-GMDN 208
QY 118 CIYMDANTNKTSQLDNGMRDQ---MSFLMS-----KVGS--FLAVGYVGNLXIYNHQ 167
DB 209 TIRLMPKSGQC--LGDALRGHSMKWTSLMEPIHLVKKPSKPRLASSSKDGTIKIMDTV 266
QY 168 TSKRIPIVLGKHKRTKTCGCNNAENLXALGGEKMITV--SNEGDTIRTOTVR----- 218
DB 267 SRVCQYTMGHTNSVSCVKWGGCLLYSGSHDRVTVMIDINSOGRCINILKSHAHVYVNL 326
QY 219 -----SEPXNMQFLM-----KMDRTSAESMISVVLGKRTLEFL 254
DB 327 SLSTDYALRIGAFDHTGKKPSTPEAOKKALENYEKICKKNGNSEEMVATASDDYTFLW 386
QY 255 NLNEPDNP-ADLEFODDFGNIVCYNMYGDRIMIGFCGHFVVISTHTGELGOEIFOARN 313
DB 387 NPLKSTKPIARMGCHQKLVNHVAFS--PDGRYIVSASFON--SIKIMDGRDKKFIPTFRG 442
QY 314 HKDNLTSIAVSQTLNKVATCG-DNCIKIODL-----VDLKDMVILNLDEENKGLGTL 366
DB 443 HIASVYQVAMSSDCRLVSCSKDPTLKVMDVRYTRKLSVDLPGLIKTKLYVD----- 492
QY 367 WTDGQ 372
DB 493 WSYDGK 498

Search completed: May 21, 2002, 14:50:30
Job time: 50 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 21, 2002, 14:49:40 ; Search time 31.32 Seconds
(without alignments)
4239.957 Million cell updates/sec

Title: US-09-729-653-2

Perfect score: 7257

Sequence: 1 HSLIGRCSRLGDNNAVAC.....AAQLKRIISDCTQYLRTHEEL 1382

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1157	15.9	519	2 T27880	hypothetical prote
2	924	12.7	519	2 T27881	hypothetical prote
3	270.5	3.7	1758	2 T34393	hypothetical prote
4	263	3.6	1198	2 T43484	hypothetical prote
5	251.5	3.5	1462	2 T00345	hypothetical prote
6	251.5	3.5	1653	2 T14758	hypothetical prote
7	209.5	2.9	1307	2 T19497	hypothetical prote
8	190.5	2.6	1119	2 T15842	hypothetical prote
9	187	2.6	1124	2 H88772	hypothetical prote
10	182.5	2.5	1653	2 A36349	protein F23B2.4 (1
11	178	2.5	1047	2 T21306	clathrin heavy cha
12	177	2.4	1349	2 S51471	hypothetical prote
13	161.5	2.2	558	2 T17224	killer toxin insen
14	161.5	2.2	1223	2 T17345	hypothetical prote
15	161	2.2	1666	2 T38393	hypothetical prote
16	160.5	2.2	421	2 T40614	clathrin heavy cha
17	158.5	2.2	920	2 C96831	G beta repeat prot
18	157	2.2	702	2 T27730	hypothetical prote
19	155	2.1	545	2 B70374	hypothetical prote
20	155	2.1	590	2 D83063	conserved hypotet
21	155	2.1	1455	2 E75199	hypothetical prote
22	155	2.1	1787	2 G97222	DNA-directed DNA p
23	154.5	2.1	576	2 T49741	hypothetical prote
24	153.5	2.1	618	2 T49741	hypothetical prote
25	152.5	2.1	926	2 T39563	related to stress
26	152.5	2.1	2386	2 T39911	probable coatmer
27	151.5	2.1	1253	2 T40302	rad3 checkpoint pr
28	149	2.1	561	2 T41301	hypothetical prote
29	148	2.0	1258	2 A12155	probable signal re
					WD-repeat protein

30	148	2.0	1557	2 T18412	lipid-binding prot
31	147	2.0	378	2 E96950	TPR-repeat-contain
32	146.5	2.0	1708	2 AE1866	WD-40 repeat prote
33	146	2.0	905	2 T38314	probable vacuolar
34	145.5	2.0	2948	2 T22664	hypothetical prote
35	145	2.0	1747	2 AC1842	WD-40 repeat prote
36	144	2.0	1189	2 AH2154	WD-repeat prote
37	144	2.0	1556	2 S59393	probable membrane
38	144	2.0	1683	2 T30885	Clathrin heavy cha
39	143.5	2.0	1681	2 S42369	complement compone
40	143	2.0	1711	2 AD1842	WD-40 repeat prote
41	142.5	2.0	600	2 T23497	hypothetical prote
42	141.5	1.9	1675	1 LRRTM	clathrin heavy cha
43	141	1.9	646	2 AE2428	hypothetical prote
44	140	1.9	2875	1 RRYUTM	genome polypeptide
45	138.5	1.9	1009	2 AF2033	hypothetical prote

ALIGNMENTS

Query Match	Score	DB 2	Length	519
Best Local Similarity	45.0%	Pred. No. 3	3e-66	
Matches	230	Conservative	114	Mismatches 157, Indels 10, Gaps 4
QY 724	ACLNHEVEFAIRVYRRIGNVGIVMSLEQIKGIEDYNLLAGHLAMFTNDYNLAQDLYLAS	783		
DB 2	AALLDSVGMATKIFREIGDAMVFTALELETIEEKNLLHAQIYTLISRYDDAEQLYLES	61		
QY 784	SCRPALEMRRLDQWDSALQAKHLAPDQIPRISKEVYALQLEFADYNNALAHYKGT-842			
DB 62	SRMEALNMRDLLEMPKALVLAETMNPKEIPLSKYQOELTLTGDHANSNLYYKGVW	121		
QY 843	-----TGDNKHEDEACLAGAAGVMSIRMGDIRGVNVALKHPKRVLRKDCGAILNNKOPS	897		
DB 122	ENPQNPPELQENHEICGOSIARMAIKTGDRLKRVQALAKOLEGIVNVRDCAIILEQMKQYT	181		
QY 898	EAQQLYEKGLVYDKASVYIRSKNMKAVGDLPHVSSPKIHLQYAKAKGADGKRYEAVVA	957		
DB 182	EAQQLYEKGLVYDKASVYIRSKNMKAVGDLPHVSSPKIHLQYAKAKGADGKRYEAVVA	957		
QY 958	YENAKQWQSVIRYLDHLNPNPEKAVIVVETOSLDKAKVAREFLDGYGSAIOFLVMS	1017		
DB 242	YENGRQYDQVQVRLDPLNDPDAVVRVRSRSIEGAKVAREFLDGYGSAIOFLVMS	301		
QY 1018	KCNNEAFTLAQOHKWEIYADIIGSEDTNEDQSALEFEGEKRYLQAGKFFILCGOYS	1077		
DB 302	QCVQGEFELAEKKNVAREAKALEQHGNSQALE-LAEYNNRNDKMFMAKFTTQAGQGN	360		
QY 1078	RAKHEFLKCPSSDVAIMEALETVGQARDELITNQLIDHLLGENDGMKDKAYFLRYM	1137		

```
Df      361 NAINLFFR - NOGDENCVALLAVDGGIKSKDKTLNNKLVKFLLGE - DGNVADPQQLFRLYV   417
OY      1138 ALKQYLEAAQTPIITIAAREOSAGNATRNNAHDVFESMAYELKSÖRKIPDESEMATNMILHST    1197
            I : ::::|::::: | : |||| : || | : :: : : :: | : :: | :: | ::
Db       418 GGFRTCAAATAVVVAQHIOAKGKNRIARLDLFPMOHOLFREKKMRIRIPLDMNKSLMAIHST        477
OY      1198 ILTVKHIVAKNGDHMKGARMLRVANNISKEPS 1228
            |:|| : : : : ||::|| | : ||| : 
Db       478 IIIVKALINRRKETFLLAARILLIFRCGIORPT 508

RESULT          2
T27881
hypothetical protein ZK520.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:date: 15-Oct-1999 #sequence_revlsion 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T27881
R:Steward, C.
submitted to the EMBL Data Library, March 1997
A:Reference number: Z20434
A:Accession: T27881
A>Status: preliminary; translated from GB/EMBL/DDBU
A:Molecule type: DNA
A:Residues: 1-634 <WII>
A:Cross-references: EMBL.Z93822; PIDB:CAE07301.1; GSPPDB:GN00021; CESP:ZK520.3
C:Experimental source: clone ZK520
C:Genetics:
A:Gene: CESP:ZK520.3
A:Map position: 3
A:Intons: 53/2; 75/2; 120/2; 187/1; 318/2; 341/1; 455/1; 530/3
```

	Query Match	12.7%	Score 924	DB 2:	Length 634	
	Best Local Similarity	30.4%	Pred 3.8e-51			
	Matches 204	Conservative 130	Misclatches 239	Indels 98	Gaps	8
Qy	56	PIQFAWQKTSQNYLAVTGADYIVKIFEDRRGOKRSEINLPGNCVAMDMDKDGVLAIVAEK	115			
Db	37	PIIHRH-RENGHTVAVACANNFTYIYDKKGNVIDALNPTGKILDIAMDREGVLAIVAN	95			
Qy	116	SSCIYLMNANTKTSOLDNG---MRQMSFLLMSKYSGLANGVYKGNLIYNNQTSRKI	172			
Db	96	TGTITLMDVNSRNTDVEEGATSSKSLPPLCLAMSPSPPTLVGNNAAGNTVYNNHRTSRI	155			
Qy	173	PVLGKHKTRKITCGCMAENMLALGEGDKMITVSNQGDITROTQVRSSEPNNOFLMKKD	232			
Db	156	AVMGKHQSRVTOITQIYPEIYVIVCSDDNLTSLTLEGTIVSTTTTGEPFNMDY--GSVN	213			
Qy	233	DRTSAAESMI SYVLGKRTLFNLNEPDPADLEFQODFGNITVCYNNWGDGRIMTGFSGC	292			
Db	214	GKGSGVTVMSVYVIGKIKILMLAHYNNLADPNNVLOFEKYNHISYRWFNDYIILGFDRG	273			
Qy	293	HEVNVISTHIGELGQELFFQARNHKDNLITIAVSQTLNKKVATCGDNCIKIDDYDLADMYI	352			
Db	274	YLLISISAHNNEIGSELVSFLEYRGYLAISIAVSTSFENKLTITIDNMKRVBDLDELTVTML	333			
Qy	353	INLDEENKIGLITLMDDOLLASTQROSLNHYETLKLPTLGDACSTRJIAVLSLLEVY	412			
Db	334	TEIEIE-KULSEIYEEDQIVAVSSQSVLSIFVYKMTFLAASYNNISICITLNLQIVY	392			
Qy	413	ANPVEGELPIVSVDEPNFVAVGLYHLAVGNRRAFY-----VIGENAV	458			
Db	393	VAVEIEKGSSTLELNIPTVMGIGLPNLIAVANNNTVEFYDHTPAOMQAALQSTQSA	452			
Qy	453	KR-----LDMETLGVASISICLSDYDAALFBEKVQDLHLESILDAQERETRLRP	510			
Db	453	EKPITVAEPIHARVELSLTVTNIQILMYMAAANFGSRRLHRIIRN---SEDNVSIIEP	507			
Qy	511	AVDDCKRIILHARVELSELTFTDGVQVQFYIEDMQFVNDYRHPVSVKIKIFDPNCTRLV	570			
Db	508	EANRRATLYSYALTIERFL-----	526			
Qy	571	FIDEKSDGFVQCPVNDATYEIPDFSEITKIGVLWENMPMDKGVFIAYDDKVVYTVYVHKDT	630			

[illegible]

```

Query Match 3.7%: Score 270.5: DB 2: Length 1758;
Best Local Similarity 18.7%: Pred. No. 1.3e+08;
Matches 270; Conservative 220; Mismatches 493; Indels 463; Gaps 68;

65 SCNYLAVTGADYIVAFIFRHGQKREINLPGVCAMDMDKDDVLVAIAEKSSCIYLDMA 124
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
27 NSRRAIACSDRSVALLDENGVOKORF-----TKRPF-----AKYCKSFYVLC--- 71

125 NNNTKSQLDNGMRDMSFLMSKVSFLAVGTGKGLXLYNQTS--RKILPLGKHTK-- 180
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
72 -----MTFSPDSSRIALIGSDNVLFLYKGTGSMNEKKVLYNKFVOPS 113

181 RITCGGMAENLXALGEDKMTIVSNQEGDTIRQTVRESEPYXNMQFLLKKMDRTSAAS 240
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
114 AVTCLSWPDDKLIYQGLDKVRIG-----LTKNCCS-----LYKTD-----ET 154

241 MISVYLGGKTTLEFLNLNEPDNADLEFQDFGNIVCYNMYGGGRIMIGFC----- 291
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
155 VVSIGTHPKRTSFSVAH-----QD--GSIIILNFSSTQSKI--CTLOVPYML 199

292 -----GHFVVIS-----THTGELQELFFQARNKNDLNTSLAVSQTLYNKV--ATCG--- 334

200 VETNGLVAVTSDRRLVSLYTGENVVQQQCDYNDQSEKFFSSISCDPTAONVGVTCGGVF 259
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
335 --DNCIK-----IOLADV-----LKDMYV-- 351
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
260 SVDCLLRKMLKSPRETTYVAVSHVILRVTNDTRTVISNKGALIDELIKGKRYVIG 319

352 ----TLNDEENKGLTSLWTDGQALLASTQKSSLAHFLTKPLDAGSTRVAYLTS 406
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
320 YTSSTIITADTESQRFSELEM-----QSGGHEKEFYD--FNMC-----LIIN 360

407 LLEVYVAN--PVGELPRTIVSVDEENFAVGL--YHLAVGNNNNAWPLVGENNAVKKK 462
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
361 AGEVTVYVEYGVDGSLGWRTSLTSHLLSVQLKAFKFSKSYGNRFPD-----ESSVSYSE 414
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
463 DMEVLTGVAISLHSDYAAALFEKGVQHLHLEISLIDAOEERETRLFPVADDKCRT---- 518
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
415 HSDL--NGQICVFSDFL-----QRTTSVYFLIKAKAIFNDL 451

519 -LCHALLSDLLYGTDTGVGVYFIEDMQFVNDYRHHVSVKILFPDPNGNTRLVFIDEKSD 577

```

```

Db 452 QLEPHVNPVLLINGQOESFINHTGAVDMIELNE-----RASKLLYRDKRS- 496
QY 578 GFVYCVNDAITYEIPESPFIKGVL-----WEMPMRDKGVFIAYDDDKYTYTFPHKD--- 629
Db 497 -----KVLTVLISSDQSRVLLSFCTYQVWPMPSDVIYQSGDNLSIW-YNPLPE 545
QY 630 -----TIQC-AKYLIGSTFVPPAHKPLLLYNGELTCQOTQSGKV-----NNIYLSHGFLS 679
Db 546 QVTNMIKIGVEAVLADADPT-----EVIQGEPTAVAYVELDNTQIE---FGA 590
QY 680 NLKDXGPDELPRPLAHNLMLKRPDSAMEMCRILNDEAANMELARACLIHMEVEFAIRVYR 739
Db 591 ALEKRPEDRAVAFLESN---TSGDAYSM-----MIRVEMALEHGNLFVAQCVCYA 638
QY 740 RIGNVQIYWSLEQIKGED-----YVLLAGHLAMFNQVNLADLYLASSCPI 767
Db 639 AINDVAKVRRLHDLITADASISIGDGTHTFYKVRAMLAIMGRKFEARIFLEQNDTE 698
QY 788 AALEMRDLOHMSALQALAKHLAPDOIPIFSKEY-----AIOLEFA-GDYVNA 834
Db 699 SALGWTSLHKMDEALELAKVLYNPREYEQULTYLRALSTGDSKRAELKVSDDGTLSA 758
QY 835 LAHYEK-----GITGDNKE---HDEACIAGYAOMSIRMGDIRGVNQLKHPSRVYLR 884
Db 759 IQLYIKSNKPLSALSAANDSVLSQDENILRQIADSLVK-----SOLYD 802
QY 885 DCGAILIENKQSEAAQLYIKGLYDKA-----ASYIRSKNWKAKVGDLLPHVSP 935
Db 803 KADVYEKLEKDFKAVEYFKKGDAYGKAIQALAFAPPEKVLTLEQENG----- 850
QY 936 KIHLOYAKAKEADGRY-----KEAVVAYENKOKOMOSVIRITLDHLNPEKAVNIYRET 988
Db 851 -LLEIYIGQVDAVNVHFEVNEANDLKAVEAIIKKEW-----PKALSTY--- 892
QY 989 QSLDGAKM-----VAREFLOLGDYGSALQFLYMSKCNNEAFTLAQOHNK---MEIYA 1037
Db 893 ENIQDQKVRITGYGEIADHYSNKGDFERABRLFEVAGLFNDALIMYGKNNKWIADPRLSE 952
QY 1038 DIIGSEDTNEDYQSTALTYEGEKRYIQAGKFFLLCGOYSRALKHPLKCSSESDNAIEM 1097
Db 953 EFGGRATIS-SYLAKAEDLDEHGRFAEAEOLYITGMIPKHAIOY----- 997
QY 1098 AIFETGOAKDELTLNOLIDLGENDMPRKADAYLFRLYMALKOYREAAQTATIIAREEQ 1157
Db 998 --DRVGRDDVL---RLVEKYHGEH--MHETKRKF-----ATQYERGD---LKAABEQ 1041
QY 1158 --SAGNYRNAHDVLFYSY--AELKSOKIKI-PSEMATNL--MILHSY-----ILVKIH 1203
Db 1042 FLKAGDFRSA---VNNYKSEKMSMDAYRIAKTEGEMEKQVLFEMAKSIGDDAAVKLL 1097
QY 1204 VKRGDHMKG-----ARMILRVANNISKPSFHIYPIILTSYIECHRAGLKKSASF 1251
Db 1098 NKHGMIEGIDFACETGAFDLAFDLARIQ--AKDRMGTAVHRLAQILE-EGRELEASK 1153
QY 1252 SFPAMLMRPEY-----RSKIDAYKKKIEGVMRRPDISEIEAATTPCPC 1296
Db 1154 HYVEAIKLNLTNTTWCQAVPSRFDLQKRLGKPKPELLAVEFPIRDNMDADAEVAK--DHC 1211
QY 1297 KFLUPE 1302
Db 1212 ESTLPD 1217

```

RESULT 4
 T43484
 hypothetical protein DKFZp434K016.1 - human (fragment)
 C:Species: Homo sapiens (man)
 C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jan-2000
 C:Accession: T43484
 R:Koehner, K.; Beyer, A.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
 submitted to the Protein Sequence Database, December 1999
 A:Reference number: Z22514

```

A:Accession: T43484
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1198 <AA>
A:Cross-references: EMBL:AL133565
A:Experimental source: adult testis; clone DKFZp434K016
C:Genetics:
A>Note: DKFZp434K016.1

Query Match      3.6%; Score 263; DB 2; Length 1198;
Best Local Similarity 18.2%; Pred. No. 2.2e-08;
Matches 245; Conservative 197; Mismatches 506; Indels 396; Gaps 54;

QY 116 SSCITVMDANTNKTSQLDNCRDMRQMSFLMSKVSFLAVTGVGNLXITHYHQRSRKPV- 174
Db 70 SSDFGLSMSPQKSVSK--HKSSSKIIICCSWTDNGQYALALMPENGILSTIRKNKEBEVKIE 127
QY 175 -LGKHTKRITPCGW-----NAENLXALGEDKMI-----TVSNOEG 209
Db 128 RFGSLSPIWISICMNPSSRWESFMNRENEDA---EDVIYRIYQIETPSLKSAVAYSISOG 184
QY 210 DTRQTVQVSEPYXNMQFELKMDRTSAESMISVY-LGKKTLEF-LINLEPNPADLEF 267
Db 185 SEAESEEPEDSDSPR-----DNLLEERNIDILAVADMCKVSEFYQLSGKQICKDRALNF 238
QY 268 QDPGNIVTCNMYGDGR-IMIGFSCGHFYVISTHTGELGQEIFQARHNKNLTSIANSQT 326
Db 239 DP-----CCISTYTKGEYILG---GSDKQVSLFT-----KGV----- 269
QY 327 LNKVATCGDNCIKIQDLVDLKDQVYIILNDEENKGLQTLSTMDGOLLALSTQSLHVF 386
Db 270 --RLGTGV-----EQMSWWTQCAKSDSNVYVVGCGDDGTISFY 305
QY 387 LTKLPIIGDACSRIAYLISLEVYANPVYEGELPTVSVDPENPVAVGLYHLAVGMN 446
Db 306 QLIFSTVHGILKDRYAVRDSMTDVIYOHLLIT-EQVKRIKCELVKIAIYRNRLADOLPE 364
QY 447 RANFYVLGENAVKVKLDMETLGTVASICLSQVAAALFEKQVQHLIESLIDAOEER-E 505
Db 365 KILYELYS--EDLSMNY-----RVKEKTIKKECNL-LVVCANHIIICOERLQ 412
QY 506 TRLPFAVDK-----CRILCHALTSDFLIYGTGTGVQYFIEDQFVNDYRHPV 555
Db 413 CLSPSGVKEREMQMESLIRIKYIGRPGREGILLVGLKNOQILIKIPVDNLFATYLLKQAT 472
QY 556 SVKKIPEPDNGTRLVFIDEKSDGFVYCPVNDATYEIIPDSPTIKGLVEMPMRDKGVFI- 614
Db 473 AVRCIDMSASRKKLAVVDENDTCLVY---DIDFKELLFQEPNANSVAMNTOCEDMLCFSG 529
QY 615 -AYDDDKVYTVYVPHKDTIQCAKYLIGSTKVPFAHKLPLLYNG-ELTCQOTQSGKVNNIYL 672
Db 530 GGYLNTKASTFVPVHROKLG--FVVG-----YNSKILFC----- 561
QY 673 STHGFLSNLKDQGPDELPRPLAHNLMLKRPDSAMEMCRILNDEAANMELARACLIHMEVE 732
Db 562 -LHVF--SISAVEVPOSPARVYQ-LDRKLFKFAVQIACLVGTDMDLAMEALEGIDFE 617
QY 733 FAIRVYRRIGN---VGIVMSLEQIK--GIEDYNLLAGHLAMFNQVNLADLYLASSCPI 767
Db 618 TARKAFIRYODLYRLLELISIEERKKRGFTNNDLFLADVSYOGKFHEAKLYKRSCHEN 677
QY 788 AALEMRDLOHMSALQALAKHLAPDOIPIFSKEAIOLEFAGDYVNNLHAYEKIGTGDNK 847
Db 678 LALEMTYDL-----CMFETAKDPLG-----SGDPK 702
QY 848 EHDDEACIAGYAOMSIRMGDIRGVNQLKHPSRVLRKDCGAILIENKQSEAAQLYEKGIL 907
Db 703 E-----TKMLITQADMAR-----NIKEPAAVAYEMVYISAG 732
QY 908 YNDKASVYTRSKNMAKVGDLLEPHVSSPKTHLOYAKAKKEDGKRYEAVVAYENAKKQOSV 967
Db 733 EHVKAIEI-CGDHGW---DMLIDI-----ARKLDKAREPFL-----L 767

```

QY 968 IRIYIDLNNPEKAVNIIVETOSLDGAKKVARFLOLDGYSALIOFLVMSKCNNEAFTLA 1027
 DB 766 CAPYLKLLKSDPGVA-----AETYLKMGKLSLYOVLHETQKNDENFAFG 811
 QY 1028 IQQNMKEIYADIIIGSDTNNEDYQSIALLYEGEKRYLQAKFFLLCGYSRAKHEFKCP 1087
 DB 812 EKHPEFK-----DDIYMPAOWLAENDREFEAKAFHKAGROREAV----- 852
 QY 1088 SSEDVAIEMAIETVGAKDELITNOLIDHLLGENDMGKDAKYFLRLMALKQYREAO 1147
 DB 853 -----QVLEQLNNNAVASR-----NDAAVY--WMLSMQCLDIAO 887
 QY 1148 TAILAREDSAGNRYNADVLFMSYAEELSKOKIKIPSEMATNMLHSYILVKKIVKNG 1207
 DB 888 DP---AQKDTMLGKFYH-----FQRLAEL-----YHGVAH1HRTEDP 922
 QY 1208 DHMKARMLIRAVANNISKPPSHIVPLISTVIECHR---AGLKNSAFSPAAMLMPREYS 1264
 DB 923 FSVHRETELF---NISRELSHLPKDTPSGISKVKILFTLAKQSKALGAYRLARAYDK 978
 QY 1265 ---KIDAKYKKKIE---GMVRRPDISEIEEATTPCPFCFKFLPECELLCPGCKNSIPYC 1317
 DB 979 LAGLITPAFOKSIELGTLITRAKPPHDSF-----LVPLC-----YR 1016
 QY 1318 IATGRHMLKDDMTVCPCDFPALYS-----ELKIMLTESTCPM 1356
 DB 1017 CSTNPLNLNNLNCVINCROPFIFSASYDLHLVEFLYLEGITDEALISLDLEVLBRK 1076
 QY 1357 CSERLNAOLKKISDCTOYLRTBE 1380
 DB 1077 RDDR---OLEIANSSQILRLVE 1096

RESULT 5
 T00345
 Hypothetical protein KIAA0590 - human
 N:Alternate names: hypothetical protein DKFZp564L232.1
 C:Species: Homo sapiens (man)
 C:Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 21-Jul-2000
 C:Accession: T00345; T12465
 R:Nagase, T.; Ishikawa, K.; Miyajima, N.; Tanaka, A.; Kotani, H.; Nomura, N.; Ohara, O.
 DNA Res. 5, 31-39, 1998
 A:Title: Prediction of the coding sequences of unidentified human genes. IX. The complet
 A:Reference number: 214086; MUID:98290545
 A:Accession: T00345
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-1462 <NAG>
 A:Cross-references: EMBL:AB011162; NID:g3043703; PIDN:BAA25516.1; PID:g3043704
 A:Experimental source: brain
 R:Diesterhoef, A.; Lauber, J.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
 submitted to the Protein Sequence Database, June 1999
 A:Reference number: 217522
 A:Accession: T12465
 A:Molecule type: mRNA
 A:Residues: 807-1069, 'A', 1071-1462 <DUE>
 A:Cross-references: EMBL:AL080069; NID:g5262476; PIDN:CAB45696.1; PID:el491529; PID:g526
 A:Experimental source: fetal brain; clone DKFZp564L232
 A:Genetics:
 A:Note: KIAA0590; DKFZp564L232.1

Query Match 3.5%; Score 256.5; DB 2; Length 1462;
 Best Local Similarity 17.3%; Pred. No. 7.8e-08;
 Matches 247; Conservative 219; Mismatches 485; Indels 475; Gaps 64;

QY 54 GAPIOFANOQTSIGNYLA-----TGADYIVKIFDRHQ--KRSEINLPGNCVAMDMK 104
 DB 16 GSPFSTHMPVH-PFLAVAYISTSTGS---VDIYLEGECVPDTHVEHPFVAVASLCHMP 71
 QY 105 DGDVAVIAEKSSCIYLDAMNT---NKTSQLDNGM---RDMSFLMSKYGSFLANGT 156
 DB 1111

DB 72 TRILVLA-----GWETGEVTFVFNKODKEQHTMPLTHRTADITVLRKSPSGNCLLSGD 122
 QY 157 VKNLXIVNHQSRKI---PVL-----GKHTKRITGGW-----NMENI-----XALGED 199
 DB 123 RUGVLLMLRDLGRKRGVQGTPLLKHEYGKHLN---CIFRLPPGEBDLVOLAKAVSGE 178
 QY 200 KMITVSNQGDITROTQVRSBPXNMQFPLMKMD-----DRTSAEBSMLSVLAKKTLFF 253
 DB 179 KALDMFNMKSSSGSLKNGSHGELLFVSLMDGVHYHDEKGTQTVVASDSTIQMLFY 238
 QY 254 LMLNE-----PDNPAD---LEFOQDFGNIVCYNWGDGIMGFSCG 292
 DB 239 MEKRALVVTENLRSLTYTPPEKAEVKKLSGKTR-----RADIALIEG 288
 QY 293 HFVYVISTHG-----ELGCEIFOARNK-----DNLTSIAVSOTLKVATCGDN--- 336
 DB 289 SILVAVAGGAALRFMDIEGENTYIISPDKPFKEGENNCCYCKVKGLLAAGDRCRV 348
 QY 337 --CIRIODLV-----DLKDMVILNDEENKGLGTLSWTDGQLLALSTORGLSHVFLTK 389
 DB 349 AMMRKVPDFLSPGAEGRKRWALQTPTELQGNITQIQWGRKRLAVNS-----VYS 400
 QY 390 LPIILD-ACSTRIALVLSLEVTVANPVBGELPIYSDVDEBNFYAVGLYH----- 439
 DB 401 VALISERAMSSHFOOVAAMQVS-----PSLNV---CFLSTGVAHSLRTDMHIS 447
 QY 440 -----LAVGMNRRAMFYVGENAVKCLKDEYLGTVASICLSHSDYAAALFEKGVOLH 491
 DB 448 GVFATRKDAVAVANGQVAFELSGAIAISAG--TLCEPVLAMHEDENVYTESNRVQR 505
 QY 492 LIESEILDAQEERETRLPPAVDDKCRILCHALTSPLIYGTDTGVQVYFYEDMQFVNDY 551
 DB 506 TWQGTIVKQLLLEFSETEGNCFLDIC-----GNLVGTDLALFKSFDSLRSRE-AKAH 556
 QY 552 RHPVGVKKIEP-----DPNGTRLVFIDEKSDG-----FVYCPVNDATYELPDPSF 596
 DB 557 CSCRLAEVLPGVGSIARCSSGSTSTISLRKADNSPDSCICYDYVEMDTVTVFDE-- 614
 QY 597 TIKGVLMENMPMDKCVFLAYDDKYVTVYFHHDTIQGAKVILAGSTKVPFAH-----RPL 651
 DB 615 -----KTGQIDRETFISFNQE--TKNSHLFYVDEGLK-----NYYPVNHFMQSEPR 659
 QY 652 ILY-----NGELTCQTQSGK-----VNNIYLS-THGFLSNLKDQSPD 687
 DB 660 LFCVAVOETPRSPQPSANG---QPODGRAAPADVLILSFFISEHGFL--LHESPP- 712
 QY 688 ELRPLLAHNLH-----LKR----- 702
 DB 713 --RPATSHSLGMEVPIYFTTRKKPEADREDEVEPGCHHITPOMVSRRLPRDVGLEDCDK 770
 QY 703 -----SDAMECRILNDEANMELARACLHMEVEFAIRVYRIGNV 744
 DB 771 ATRDAMLHFSFVTTIGDMDENAFKSLIKLSEAVMENMARCKYQRLDVA-KVC-LGNM 827
 QY 745 GIVNSLEQIKGIEDYNLLAGLAMFTNDYNNLAQDL-YLAASCPIALEMRRDLQH----- 798
 DB 828 GHARGARALREOEOPPELEARVAVATOLGMEDAEOLYRK-----KRHDLKKFYQA 881
 QY 799 ---WDSALOLKHLAPDQIPFISKREYAIQLEFAGYVNAALHYEKGITGDKNKHEDEACLA 855
 DB 882 AGRWQDALQVAEHHDVHILRSTYHRYAGHLEASACSRSLSTYERSDT----- 929
 QY 856 GVAQMSIRMGDIRGCVNOLAKHPHSVLRKDCGAILENKOFSEAOLEYEGLYDYKASV 915
 DB 930 -----HREYVP-RMLSEDLPSL-----ELIYVK----- 951
 QY 916 YIRSKN---NAKVGDLLPHVSSPKIHLQYAKAKPADGRYKAAVAVYENAKOWOSVIRY 971
 DB 952 -MKDKTLRWKMA-----QYL---ESQEMDAALHYELARDHFSLVRIH 991
 QY 972 LDHLNPEKAVNIIVETOSLDGAKKVARFLOLDGYSALIOFLVMSKCNNEAFTLAQOHN 1031
 DB 992 C-FQGNVOKAQDIANETGNLAASYHLARQYSOEVEGQAVHYETTRAQAFKNAIRLCKEN- 1049

Oy	276	CUNYNGGRIMIGSCHGFVIVISTHNGELGQELFQARHNKQDLTJIAVSQTLNKVATCGD	335
Db	189	CLKTW-----SPTL-----	197
Oy	336	NCIOTDLVDLKDWYVILINDEENKIGTJLSMTDDQLLALSTORGSLHVFETK-----	389
Db	198	---KSKTILDEMEDEFRKLEKREVTGTAIFLN-----PPTKPREL	238
Oy	390	-----LPILGADCSRIAYLTLSLEEVYANVEGELPITVSVDPERFVAVGLYH	439
Db	239	PEHEKNETEPYQPPDRPRFVAVYARGMNO---LMRSLNDPERVYVSI---PRFKITG---	290
Oy	440	LAVMNNRANFYVLGENAVK-----KXDMETLG-----TVASICLSMD----	478
Db	291	-AKSPSPGAFALVACGNDSDKDEPTFSKIHFLSAVGIYVGPOTNDSJCITGCWESTGGLRM	349
Oy	479	---YAAALFEGKVOULH-----IESILLNOEERETFLPAY-----DDK-CRILCHALT	524
Db	350	AIAMAAGTLLIGQIRPERPKMSIETIIVYVYOKELYOYIMYDYKTDEKTKVKTHTHEN	409
Oy	525	SDF-----LIYGTGNGVOVFIETEDMOFVN---DYRHPVSXKFIPEPNGRFLVFIDEK	575
Db	410	MAFIREHCVLINRQDQSVIRHYFCQJLNSIGTSLDQ-NTTYTRPRFACVNGICATYASE-	467
Oy	576	SDGVYCPVNDATYEIPESSPTIKGYLVEMENPMKQCVFIAYODDKYVTVYFHKDTIGQAK	635
Db	468	DRYFIW-----HFVLKPFMSVOAGI---HVPEKSGDYVLEEQRTIEY-----GTK	510
Oy	636	VILAGSTRVPFAHKPRLLYNGELTCTQOSGKVNIYLTSHGFLSNLKDQPRELAPMLAH	695
Db	511	RLDSKREI-----CALCTG-----	525
Oy	696	NLMKREPSDAMEMCRILINDEAANNELARACLHHEVEFAIRYRRIGVIGMSLEQIKG	755
Db	526	-----DFFPMALLSG-----GIRYSLNDGYI-----	548
Oy	756	IEDYNLLAGHLAMFTNDYNIADQL-YLASSCPITALEMRDLOHWDLSQALAKHLAPQI	814
Db	549	-----TNSYVPSPINSIKINCFTPL-----AVIKIV-----EQV	579
Oy	815	PEISKEVAIOLEFAGOVVNMALHYEKGITGDNKEHDEACLAGVAQMSIRMGIRRGVNO-	873
Db	580	PEOFLLY-----EPFGDELKLI-----YTSDKRD-----VWQYEMDQNNIS	614
Oy	874	---ALKHPSRYLKRDCALILENNKQFSEAOLEYEKGLYDYKKAASYIRSKNNAKVGDLLP	930
Db	615	NMLALTKOKRILICDGSISLE-QSSVNGSILFLQ-----NLVYRANIEKI---LLT	662
Oy	931	HVSSPKIHLQYAKAKREADGRYKRAVYAYENAKOMQSVIRIYLDHILNPEKAVNIYRETOS	990
Db	663	PENFTKTCIVEVMIKA-----KQDVNMVLVNAKML-----EADIDYAEKSPH	703
Oy	991	LDGAKNVARF-----FLOLGQGSALQFLV-----	1016
Db	704	SELSMSTIANAVEFKHAFCDEAHAFVKIGDV-AGVQDLAVEMYKKVGDYASVYDLIRNDD	762
Oy	1017	---SKCNNEAF-TLAAQHNKEIYADIIGS-----EDPTNE-----DYOSIATYE	1058
Db	763	DEGCG--DAPRNAHEHYEMEMEBEAKAYASYCGOTRANQIDCLINGNLFGELYAVLARLS	820
Oy	1059	GKKRYLOA-GKFFLLGQVSRALKHFLKCPSSDVAIEMALETYQAKDELITQOLIDH	1117
Db	821	DQSEFMEMGCAFTSRGCMQDAVECYLR--RSLPQKALHACKELNOMQKAPFIAD--ANH	876
Oy	1118	LLGENDQMPRDAKTLFLYLAALQOYREAOATAIIRREGOSAGNFR--NANDVLESMYAE	1175
Db	877	M-----ENIADKRKRYVPEELKQCYVLG--AALVENHROTIKELQKIDKHNFLDALDD	930
Oy	1176	---LKSOKIKIPSEMAFNILMILSHYLVKITHVNGCHMGKARMLIRVANINISKF--PSHI	1230
Db	931	ESGLASQDSRILENTMNGALAFHFHMLMAOQHFE-KRIDALQTLISLSDYEFLDPRAEV	989
Oy	1231	VPILTSTVIECHRAGLKNSAFSFAAMLIMREYRSKIDAKYKKKIEGMWR-----PDISE	1285

```

Db      990 HSM1LAANANQFETCSKA-----MMRLFEFEEDDAEKDEEMNL5FRLPSENPNP 1043
QY      1286 IEEATPPCFKFLPECELLCPGCKNSIPCIATGRHMLDDMTVCHCDFPALYSELK 1345
Db      1044 -NSAVACSVCDARKIDPDLQCECQTKFPFCIASGRLLDNIFFWLCRCRKHRAHQHRI 1102
QY      1346 IMLNTESTCPMCSE 1359
Db      1103 ---KYNCCPLCHD 1112

RESULT      9
H88772
Protein F23B2.4 [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
C:Accession: H88772
R:anonymous; The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating bio
A:Reference number: A75000; MUID:99069613; PMID:9851916
A:Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/projects/C_e
A:Accession: H88772
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1124 <STO>
A:Cross-references: GB:chr_IV; PID:CAB05177.1; PID:g3876274; GSPDB:GN00022; CESP:F232
A:Gene: F23B2.4
A:Map position: 4

Query Match      2.6%; Score 187; DB 2; Length 1124;
Best Local Similarity 18.9%; Pred. No. 0.0015;
Matches 198; Conservative 145; Mismatches 395; Indels 310; Gaps 45;

QY      76 YIVKIFDRHGKRSSEINLPNCV-AMDMDDQDVLAVIAEKSSCIYLDAMTNTKTSQLDN 134
Db      81 WVDKILD-----ENNKGVCITYDLAKRPSOSELLAADNK--VYLPDVEGGQMOTLK 131
QY      135 GMRDQMSFLMSKVSFLAVGTGKGNLYN--HOTSRIPIVGLKHTKRI----- 182
Db      132 GHKDLVYVAVMSHNGELFASGGADKLVIIMNEKHGTLRY-----SHTDVIQCMENPCNQ 187
QY      183 -----TCGC-WNAE-NLXALGGEDEKITV-----SNQ 207
Db      188 ILLTALNEFGLMSTADKNVIKORSVVRCCSCAMWTDGTIFAIGHDSTITLIRKTKNTATE 247
QY      208 EGDTRKOTQVRSE-----XNMQFLMKMDRPTSAESMSIVYLKGTTFEFLNLN--- 257
Db      248 EESIIIQDR--NEPIWGIAFSSNRFFASRDQSGNMGIDELMAVLDNMKKTLSFYLDGTF 305
QY      258 -EPDNPADLEFQODEGNIVCYNWYGDGRIMIGFSGCHFVIVSTHTGELGOELFOARNHKD 316
Db      306 IESKN--LEFEPH-----CISYCLNGEYLLIGSGDKLIKTYRGRGVLGTVAAQDHWIM 357
QY      317 NLTSAVSQTLKVAATCGDNCIKIDLDVLDKMYIILNDEENKGLGLSWTDGQGLAL 376
Db      358 SYTVRPNSTQV--AMGCVDTGTACYNLV-----FSTVHCVDHAR---Y 395
QY      377 STQGRSLNHFLEFLKLP--ILGDACSTRIVAYFLLEEVVANPEGELPI----- 422
Db      396 ANRKSMTDVFYONLEKRRSSNICHDLYKKMSLYDTKLAVALQSDIKIYKOTGVSKNER 455
QY      423 -----TVSDVEPNFAVAGLYHLAV-----GMNRRAMFYVLENAVK 460
Db      456 RKQLYTLQDTIRKDLSESLMAYVTHGHLYVNCDEKTECYDFGKIKRSM-----N 505
QY      461 LKD-MEYLGTVASTLHSYAAALFEGKVQLHLESELIDAOEEETRLPAVDDKRI 519
Db      506 MKSIVRYL-----RVLGPAHRET----- 524

```

Query Match 2.5%: Score 182.5; DB 2; Length 1653;
Best Local Similarity 18.4%; Pred. No. 0.0052;
Matches 288; Conservative 222; Mismatches 524; Indels 535; Gaps 75;

12 LGDGNNAVCIVETGCACVLHSSRSGSSVE---MKRIFSL-----LEKWTGAPIOFAWQ 62

Db	54	LAKGEVTRKMMGDSALMHPSONAVISRANGTIVQJFNETSKLSFLLDEPVLF	WR	112
Qy	63	KTSNGYLAVTGADYIV-KIFDRHQKRSEINLPBNCVAMDMDKDGDLVIAIEKSCIIY		120
Db	113	WLSETTLEGFVARSILTSNED	GN	145
Qy	121	LMDATNNTSOLDNCRDQMSFLL-SVQGSFLAVGTGVKGLXLYNNQTSRKPIVGLKHT		179
Db	146	LHNALNNT-QILINVAKN-LDMFAVGLDENGHIAGLDQLFSKO-RNIS		194
Qy	180	KRITCGMNAENLXALGEDMITYVSNQEGDTIQTQVREBPNXMOFLKMDRTSAE		239
Db	195	QADIGHVAIFINILLEGSGTPVOY	FVTGNRAATTGAGE	233
Qy	240	SMISIV	LGKKT--LEFLNLNEDNADLEPOODGNGICYMMYGDGRMIG	288
Db	234	LRIEIDHDASLPYOXQETTDIEFFPRDPTDPIAVQSEKGIYLLRKYG		286
Qy	289	FSCGHFVISTHTEGLOEIPQAR	NHKDL	321
Db	287	PHLYELETGNLNVNRTAESVTAPARYHENGSIACINKKGOVLAVEISTQ		339
Qy	322	AVSOTLNK	VATCGDCIKIQDLVDIKMVIILNDEEN	KG 361
Db	340	IVPYLLINKLSVVALALIVATRG-GLRPADDLFQOFESLILNDQYNAKVAASSTLRN		398
Qy	362	LGTLSMTDGO	LTAIST-QRGSCH-VFLTKRPIIGD	395
Db	399	QNTIRLKNIQAPREASIPILLYFTSTLLDKKLKEETTELAR-PVLOQDRKOLFEXLK		457
Qy	396	ACSTRIAVYLTSLLEVTA	NPVEGELPIYVSVDPERN	431
Db	458	EDKLECSSELDIQVAPRTTALACLYRAGAHAKVISCLELDQOFKTIIPYCKQGYQRP		517
Qy	432	FVAQVLLYLAVGMMNRAPFYLYG	ENM-VAKLDM-EYLQTVASICLHSD	478
Db	518	FLVYL-TSSLRSPRBAEFAVSLLQNEPTASQIDIEKIADLPFSPONHIQOQTSLLDAL		576
Qy	479	YAAALFEKRVOLHLESEILDAOE	ERETRLF	509
Db	577	KGDRPDQGHLOTRVLEVLLHAPVOALIGNNIFSHYDKPTIASISEKAGLYORALEVN		636
Qy	510	PAVVD-KCRILCHALNDEL-IQDTGVQYQ	YIEDQFQVA	549
Db	637	TDIKIKICCVNHTALPDLMLVGFGLKLNQEOSLACSKALMDNNIQANIQTVOVATKFS		656
Qy	550	DYRHHVSVAKTIFPBPNGR	LVFDEKSDGFVCPVNDATAYEIPRDSPTIKGV	601
Db	697	DLIGSTLILKLFEDYQNAEGLYUULASLVNTEKDX-VVYKYL-EAANAKKQYR-ELERI		753
Qy	602	LMENRPMDKGVFIAYDDDKVITYVEHKD-TIOGAKVILAGSTKVPFAHKL-LYNGELT		659
Db	754	YCDN	NVYDERPVKNFL-KDANLEDDLPVIVICDFEFHEMIILYLSQ	801
Qy	660	QOTQSGKNNNLYLSTPHGLSLMKXGDELRLPMLAHNLMLKRSDSAMEMORILLDEAAM		719
Db	802	NLKFIETY	VQOVNSKTAQVVGALLDMD	C-DEAFIO 836
Qy	720	ELARACHLHMEV-EFALIVYRRIINVGIVNSLEO-IKIGIED	YNLLA	763
Db	837	SILQSVLGOVYINELTTFVEKRNKLKILPRLBESLSQIGQDQAVYALAKIITYDSNPS		896
Qy	764	GH	LAMFTNDYNL-AODLYASS	784
Db	897	EKFLEKENDQYDTLVDVGNCEKREPYLAVIAYEKGQONDDILRTINENSMKYQARYL		953
Qy	785	CPIALLEMRRIQDQNDAL	OLAKHLARPQDIPFISKEVATQLEFADQVNAL	835
Db	954	LE-RSDLDLNNKYLQNDENIHRKQLDLSVTSVGPILTDEPVPSTLVQAFPMNGLK		100
Qy	836	AHYEKGITGNKEHDEACLAGVAAOMSTIRMGDIRGVNOALKH-PSFVLKRDGCATL		890

Dh 1008 LELIELEKILBSPSPENVALOGLLLS-----AIKYEPTKV-----SSYL 1050
Qy 891 ENMKOF-----SEAOLYEGKLYYDKAASYIRS-KMKAVGDLPHVSS 934
Dh 1051 EKLNDVADADIAPLCIEHDKKEFAELTYDHEMGKALKVLIEDIMSLDPAASYADKINT 1110
Qy 935 PKIHLQYAKAKKADG-RYKEAVVAENAKOMQSVIRIYLDHLNPEKAVNIVRETOSLDG 993
Dh 1111 PELMSQIGTA-QLDGLRIPALIESY-----IKADPSNYEVIDIAE----- 1151
Qy 994 AKNAFFELQDYGSAIOFLVMS-----KCNNEAFTLAQOHKME-----I 1035
Dh 1152 -----OAGRYEELIPFLPMARKTLEKPEKIDGALILAYAEINKHETIENLAGSNVA 1202
Qy 1036 YADIGSEDTEDNDYOSIALYFEGEKRYLOAGFFLLCGYSRALKHFLCSPSEEDVAI 1095
Dh 1203 NLHVQDKLENEKRYKARLCYSANYSKLASTLYLGYQAADTARA---SIRKV 1258
Qy 1096 EMAIETVGAQKDELFTNQL-----IDHLLG--ENDGMPKDAKYFRLYMALQY 1142
Dh 1259 WKLVNDACIEKKEFKLAQIGLNLVHAELDELVEREESNGYFEELISLFEAGLGLERA 1318
Qy 1143 REAAQTAIITAREBQASAGNRYMAHVDLESMYAEIKS-----QKIRP-----SE 1186
Dh 1319 HMGCFTELAI-----LYSKYEPDKTFEHLKLFWSRINIRKIVIRAEQAH 1362
Qy 1187 MATNLMLHSY-----ILVKIYKNGDHHKKGARMLRVAN-----NISKFP 1227
Dh 1363 LMSLELVLAHYDEMNAALTLEKSTKLDHAYFKEVAVKSNLEIYYKAINFYKAFHP 1422
Qy 1228 SHIVPILTS 1236
Dh 1423 SLVLDLTS 1431

RESULT 11
T21306
hypothetical protein F23B2.4 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
R:McMurray, A.
Accession: T21306
submitted to the EMBL Data Library, November 1996
A:Reference number: 219403
A:Accession: T21306
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1047 <M1>
A:Cross-references: EMBL:282266; PIDN:CAB05177.2; GSPDB:GN00022; CESP:F23B2.4
C:Genetics:
A:Gene: CESP:F23B2.4
A:Map position: 4
A:introns: 16/2; 55/1; 95/2; 168/3; 266/1; 321/2; 399/3; 437/3; 591/2; 642/3; 695/3; 751/2

Query Match 2.5%; Score 178; DB 2; Length 1047;
Best Local Similarity 18.8%; Pred. No. 0.005;
Matches 191; Conservative 140; Mismatches 385; Indels 302; Gaps 43;

Qy 105 DGDVLAVIAKSSCIYLMADANTNKTSQLDNGMRDQMSFLMSKVSFLAVGTVAGNLXIV 164
Dh 27 DGSLLLAADNK--VYLFVDNEGQMQTLKGKHDVLYTVAMSHNGELFASGADKLVIYLM 84
Qy 165 N--HQISRKIPVLGKHKRRIT-----TC 184
Dh 85 NEKHEGLTRY-----SHTDVLCQMFNFCNOILLTCAINEFGMLSTADKNVIKQSVVRC 140
Qy 185 GC-WNAE-NLXALGDEKMTV-----SNOEGTIIRQTVRSR-----XNNQFLMK 230
Dh 141 SCAMNDGTTFAIGHGGGTTTLKRGKTATEEPSTIIRD--NEEWIGIASNSNRTFASRD 198
Qy 231 MDRTSAAESMISVLCGKTLFLINLN-----EPDNPADLEFQDFGNIIVCYNMVGDGRIM 286

Dh 199 SOGNPBGIDEIMAVIDMNKTLSPYSLDGFIESKN--LEFEH-----CISYCLNGEYL 250
Qy 287 IGFSCGHFVYISTHTELGCEITFOARNHKDNLTSIAVSOTLKNVATGDCNCKITODLVLD 346
Dh 251 LIGSDKILKIYTRKCVLGLTVAQMDHMTWSVYRPNQTV--AMGCVGDTICYNLY-- 306
Qy 347 KDMYVILNDEENKGGTSMWTDGDLALSTQSGIHLVFLTKLP--LIGDASTRYAVL 404
Dh 307 -----FSTYHCVDHAR--YANKRSMYDFVONLEYRTSSNICCHLVK 348
Qy 405 TSLLEVYANPVEGELPI-----TVSDVEPNFVAAGLYHLAV 442
Dh 349 MSLYDTRKLAQVSDKIQIYQOTGGVSKNERKQKLTQDTIKRDLFSLSLMVYTHGHLYV 408
Qy 443 -----GMNBNAPFVYLGENNAVKKLD-MEYLGTVASICLHSDYAAALFEGKYV 489
Dh 409 CNDEKLECYDFKGIKRSM-----NMKSIIVRL----- 436
Qy 490 LHLIESFIIDAGEERETRLFPVAVDCKRILCHALTSDFLYG--TDPIGVQYFYIEDMOFY 548
Dh 437 -----RVLGGPAHRET-----LVLGTTDGGVYKVFIDNDPIIL 469
Qy 549 NDYRHPVSVKIFPPDNGTRFLVEIDKSDGFVYCPVND-ATYEIPDSPTIKVLMENMP 607
Dh 470 LDSR-KTAIKCIDIMANRTVLAISIDP---LVCKMSDIATGFTLLQEPGCVSVENTNV 524
Qy 608 MDKGVPIANDDDKVYVYVPHKDTIQGAKVILAGSTKVPRAHKLILYNGELTCOTQSGKY 667
Dh 525 ENLFAFTNNMLHVRLAAGHTTRGVGLGVKKNKTFE--LVQYN----- 569
Qy 668 NNLYLSTHGFSLNLDKXGPELRLPMLAHNLMLKFSWAMEMCRIL-----NDEAMANEL 721
Dh 570 -----LIPLEV-PYTIHLQYIERGDFKALRIACLGYYKND--WKYL 609
Qy 722 ARACLHMEVEFAIRYRIRGN--VGIVNSLEQIK--GIEDYNLAGHLMFTNDYNLA 776
Dh 610 ANKALDLEFDVARAKAYKVRDRKMLRMWELKMKMSNGSPDAILRATILA-YTKKFREA 668
Qy 777 QDLYLASSCPITALBMRDLOHWDLSALQAKHLAPDQIPRISKEVYIOLFAGDYVNAIA 836
Dh 669 AKIFKENGFEENRAMLEFTDMRMDVDVOE-----VMTASGETKKMM 710
Qy 837 HYKGITGDKKEHDEACLAGVAQMSIRMGDIRG-----VNOALKRPSVTLKRDG 887
Dh 711 KRASWARDANOPIA-----AEMLISSGDLXKALLIIDNMLDELATISHKIDNSD-- 763
Qy 888 AILENNK-----QFSEAOLYEGKLYYDKAASVYIRSKNMA--KVGDLPHVSS 934
Dh 764 --LETMKKLSAYFLRKHEFGLASIRIQSINDMKSIYDMHVNAGHMWDFAIADRHKKYVE 821
Qy 935 PKIHLQYAKKADGRYKEAVVAENAKOMQSVIRIYLDHLNPEKAVNIVRETOSLD 992
Dh 822 -DVLYPARFLARLDRDREFAOKAFHRAGKEQEAHV-LEOLTS-----NSVNNENRPAD 872

RESULT 12
S51471
killer toxin insensitive protein 3 - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein L3502.7; protein YLR384C
C:Species: Saccharomyces cerevisiae
C>Date: 23-Feb-1995 #sequence_revision 12-May-1995 #text_change 21-Jul-2000
C:Accession: S51471; J05565
R:Du, Z.
submitted to the EMBL Data Library, December 1994
A:Description: The sequence of S. cerevisiae cosmid L3502.
A:Reference number: S51466
A:Accession: S51471
A:Molecule type: DNA
A:Residues: 1-1349 <D0U>
A:Cross-references: EMBL:U9104; MIPS:YLR384C; NID:g609423; PID:g609430
R:Yajima, H.; Tokunaga, M.; Nakayama-Murayama, A.; Hishinuma, F.
Biosci. Biotechnol. Biochem. 61, 704-709, 1997

A>Title: Characterization of IKI1 and IKI3 genes conferring pgk1 killer sensitivity on S
A:Reference number: JC5564; MIDID:97290891
A:Accession: JC5565
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-1349 <Y>A>
A:Cross-References: DDBJ:D87841; NID:g2116969; PIDN:BA20120.1; PID:g2116970
C:Genetics:
A:Gene: SGD:IKI3
A:Cross-References: MIPS:YLR384c
A:Map position: 12R
C:Function:
A:Description: confers sensitivity to killer toxin
C:Keywords: transmembrane protein
F:97-113/Domain: transmembrane #status predicted <TM>

Query Match 2.4%; Score 177; DB 2; Length 1349;
Best Local Similarity 18.2%; Pred. No. 0.0085;
Matches 245; Conservative 188; Mismatches 384; Indels 530; Gaps 66;

10 GILGD--GNAVACLVETGACVLIHSSRGSSVEMKRIFSLKTLGAPIDFAMQKTSQ 66
Db 223 GLVGNQLRPTMYMDTGTVALDSH-----ITISW-RGDC 259
QY 67 NYLAVTGADYI-----VKIEDRHQOKRSEIN-LPGNCVAMDMDKGDVLA 111
Db 260 DYPAVSVEEVPDEDEETKSIKRRARVFESREGQLDSASBPVTGMEHQLSMKPGQSLIAS 319
QY 112 IAKSSCIYIMDANTKTSQDLNMGMDQMSFLMSKVGSLAVGTAKGLNLIYNNHOTS 171
Db 320 IQRT-----DLGEEDSVIVIFFERNG-----LKHGEFDR 350
QY 172 IPIGLKHTKRTICGCWNAENMLXALGEDKMITVSNQEGDITROTQYSEPNXNQFLMK 231
Db 351 LPL-----DEKVESCVNNSN-----EALAVLANR----- 376
QY 232 DDRTSAAESMISVYLKTLFLNLNEPDNPADLEFOQDFGNIVCY-NMYGDRIMIGS 290
Db 377 -----IQLMTSKNHYMLYK-----QELVASDISYKWKHPKEDFLMTS 414
QY 291 CGHFVYISTHTGLGQ-----EIFQARNHKDNLSTIAV-SOTLKNKVTNCGNCKIKIDVD 345
Db 415 DAGFINIVDFAYKMAOGPILEP-----DNGTSLVADGGRTVITP----- 454
QY 346 LKDMYVILNDEENKGLTSLWTDGQLALSTORGLAHV-FLTKPLTIGDACS-TRIA 402
Db 455 -----LATANVPPMYRDEPRTGQNVLDVACSPSNEI 486
QY 403 YLTSLLEVYVANY-----EGELPTYSVDYEPNFVAVGLYHLAVGMNNRMVYVLE 455
Db 487 YAAINDVYLFAAVPSSIEMKKKRNHSIVCEPFKSEFT----- 524
QY 456 NAYKTLKMDMEYLG-TVAASICHSIDYAAALEGKVVQHLIESELIDAOEBERTLPFAVD 514
Db 525 SEVDSIRQVAFINDSIVGLDLD-----NLSRIALDDIQDITQPLITIVE 572
QY 515 KCRILCHALTSDF-----LIYGTDTGVV-----QYFIEDW-QFYNDYHNPYSVKIRP 562
Db 573 YDKIV-LLRSDPDYHNIYETRDGYVCOQDAGQLMETIKFPQLYRDRF-----VNR--- 623
QY 563 DPNGRTRLVFIDEKSDGFVYCVPDATYEIPDSPTIKGLVEMWPMKGVFLAYDDKYV 622
Db 624 -----VHTSAEDDD-----NMSAESSELVAFG----- 646
QY 623 TYVFKHDTIOGAVILAGST-----KYVFAH-----KPLLY----- 654
Db 647 --ITNNGKLFANOVLLASATSLSEITDSFLLTAAQHNIQFVHLNSTDFRPLVEEGVE 704
QY 655 -----NGELTCQFQSVNNIYLTSHGFLSNLKDXGP-----DELAPMLAHNLMK 701
Db 705 DERVRAIENGSLIVSIVPSKSSVYLOATRG--NLEFIYPRIVLAVER-----KNIMAKR 757

QY 702 FSDWEMCR-----ILND-----EAMNMLAR-----ACLHMEVEFAIR 736
Db 758 YKEAFIVCRTHRIINDILHDYAPELPIENLEVINQIGRVDLNLFISCL--SEDDVTKT 815
QY 737 YVRIGNVQIVWSLEDIKIEDYNLAGHLAM-----FTNDYNLAQD-LYLASSCP- 786
Db 816 KYKETLYSGISKSF-----GMEPAPLTEMQIYKKKMFDPKTSKVKVTCIAVNVLLSNE 871
QY 787 -----IAALEMRDLDHMSDALQAKHLAPDQIFPISKYAI-QLEFAGD-----YVN 833
Db 872 YKKKIYQTIITAVASQNPQNLASALKLISELSENE--EKOSCVTYLCFLQDVVWYKS 927
QY 834 ALAHYKGTIGDNKEHDEACLAGVAQMSIR-----MGDIRGVNQ 873
Db 928 ALSIVDVSLA-----LLVAQKSDMDPREYLPFLQELQDNEPLRRKFLIDYDLYGNEYK 979
QY 874 ALKHPSRVLKRCQATLENMKQFSEAOIYEXGL---YVD--KAASVYRSKNMKVGLD 928
Db 980 ALKHLSEIDKD--GNVSEVYIDVESHDLKKGALYRIDEKQVYI--NIYAK----- 1030
QY 929 LPHVSSPKIHLQYAKAKADGRYKEAVVAYENAKQWQSVIRIYLDLNNPEKAVIVRET 988
Db 1031 --HLSSNQYTDAAVAYEMLGKLEMGAVQSAKRKRREMSIAVQKF--PEEVESAEBL 1086
QY 989 QSLDGAKMYARFFLQDYGSAIQFLVMSKCNNEAFTLAQHNKKEIYADT-TGSEDTTN 1047
Db 1087 IS-----SLTFEHRYYDA-ADIOLEYLDNWK 1111
QY 1048 EDYOSIALYFEGEKRYLQAKRFFLLQGYSRALKHFLKCPSESDNVAIEMAEIYEQAKD 1107
Db 1112 E--AVALT-----CKAARIYDASVLAITKA--KRD 1136
QY 1108 ELLTNQIDILHLENDGEMPKD--AKYLFRLYMALKQYREAAQTAIITAREOSAGNYRNA 1165
Db 1137 ELL-EEVVDPGIAGEGFIATIELLADCKQINSQRLRLE-----LRAKKENP----- 1183
QY 1166 HDVLFMSYAELEKSQKTKT--PSEMAT 1189
Db 1184 ---YAFYQETEQAQDDVSVAPSETST 1206

RESULT 13
T17224
Hypothetical protein DKFp434B165.1 - human (fragments)
C:Species: Homo sapiens (man)
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T17224
R:Dueterhoeft, A.; Lauber, J.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, September 1999
A:Reference number: Z18727
A:Accession: T17224
A:Molecule type: mRNA
A:Status: preliminary
A:Residues: 1-327;328-558 <DUE>
A:Cross-References: EMBL:AL117421
A:Experimental source: adult testis; clone DKFp434B165
A>Note: the cDNA sequence contains a -1 frameshift near codon 327
C:Genetics:
A>Note: DKFp434B165.1

Query Match 2.2%; Score 161.5; DB 2; Length 558;
Best Local Similarity 22.6%; Pred. No. 0.021;
Matches 91; Conservative 47; Mismatches 148; Indels 117; Gaps 13;

QY 767 AMFTNDYNLAQDLYLASSCPIALALEMRDLDHMSDALQAKHLAPDQIFPISKY----- 821
Db 11 ALEKDFOKAEGILLRAQPRGLALNYYKAGLMSDALRCKDYVPSQLEALQEEYERAT 70
QY 822 -----ALQLEFAGDYVVALAHYKGTIGDNKEHDEACLAGVAQWQSVIRIYLDLNNPEKAVIVRET 868
Db 71 KKGARVEGFVQDARKHMEAGETSRVDCYLVKVRDSGNSGLAEKCMKMAELSTINFLLPQ 130

Qy	869	RGVQALAKHRSRLKROCGIILNEMKQSEEAOLYKGLYDKAKSVYIRSNNAKVODL	928
Dd	131	RNMEEVLA-----VGQGLIGIRKHSAAELVLTNLDYKEALDAIFEGEENK----	177
Qy	929	LPHVSSPKIHLQYAK--AKEADRYKEAAVAYENAKOMQSVIRIYLDLHNPEKAVIVR	986
Dd	178	-----KAKRYAKELDPREYDVDOH-----YKEFLKNQK-----	206
Qy	987	ETQSLDCAKKVAVR--FFLOGLDYGSAIQ-----FLVSKCKNNEAFT	1025
Dd	207	-VDSLAVGVDAIALDLYVEGGQMDKCIETATQONKYILKRYVALYATYHILREGSSAQALA	265
Qy	1026	LAQGHNR-----MEIYADLIISEDT--NEDYQSIATLFECEKRYLOAGKEFFLLCG	1074
Dd	266	LYVGHGAPAMLQNNPIYKRIFTDMVSSPGTNCQAEVHSM-----DLRVLNENLAVL	317
Qy	1075	QYSALAKHFLAKCPSEEDNVAIEAETVQOANDELLTNLDIH	1117
Dd	318	SPSSSVKTVWK--SSEAN-----SPAHEEFTYMLLIH	348

```

RESULT 14
T17345
hypothetical protein DKFZ586M1824.1 - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T17345
R:Duesterhoef, A.; Lauber, J.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, September 1999
.:Reference number: 218727

```

```
A:Status: preliminary
A:Molecule_type: mRNA
A:Residues: 1-1223 <DUE>
A:Cross-references: EMBL:AL117665
A:Experimental_source: adult uterus; clone DKFZ586M1824
C:Genetics:
C:Note: DKFZ586M1824.1
```

Query Match	2.2%	Score	161.5	DB 2	Length	1223			
Best Local Similarly	18.8%	Pred. No.	0.072						
Matches 258		Conservative	173	Mismatches	484	Indels	461	Gaps	64

```

0Y 85 GQRKSEI--NLPRGNCAMBDKXGDVLAIVAEKSSCIYIMDANTKTSOLDNGMDDQNSF 142
Db 45 GQNSHRIVFNI---CPLQDETDKOLLSTSPMRD--VKSMIDLATLECSWTLPRLSLGFAVS 99
0Y 143 LMSKV--GSEIANG-----TYKGNLXIYHNQTSKRIVLCKHTKRIKCGGMN 188
Db 100 LAFSSVDIGS-LALGVGDGMI RWMNTLSTIKNNYDVKNFMQYK-----SKVALCWN 150
0Y 189 --AANLKLAGEDEKMITVSNQEGDTIROTQVRSEXPXNQMFILMKDDRTSAEISMIVL 246
Db 151 PTKSGCLAFGTDOCKV-----GLDYTSNKPPOISSTY 183
0Y 247 GKTLFELNLNERPNADLEFQODFENICVYMWYSGDRIMIEFGSGHEFVISTHGELOQ 306
Db 184 HKKIVUTYLLAMCRRPRRPSLSGEGEDRYSALYSOGGEG-----IVLDONPWKLSG 232
0Y 307 EIFOARNHKDNLTSIAVSQTLNFKVATCSGDNKICIDVLVDLKDMLVILNDEENKGLGTLSS 366
Db 233 EAPD-----INKL-----IROTNSIKTKKLPVHT-----EIS 258
0Y 367 WTDGOLLALSTORGSIAVEFLTKRLPLGADCASTR----- 400
Db 259 WKAAQKIMALGNEGSGSEIF--QIRNKLKLTICQDNHKLVTNISMHNHENGSRDELSTYMA 316
0Y 401 -----IAVLTSLLEVTANRVEGELRTVSVDERNFVAAGLYULIADGMNNRAMFVYLG 454
Db 317 SGSNNAVIYVHNKLTVESSP-----ESVVTIT---EYRTFLSG--HDA-KITSVAN----- 362
0Y 455 ENAYVKIKLQDEUYIQTVAISICLHSD--YAAALEGKVOYLHLEISEIIDAOEERETRYLPAV 512

```

[illegible]

RESULT 15
t8393
ciatrpin heavy chain - fission yeast (*Schizosaccharomyces pombe*)
C.Species: *Schizosaccharomyces pombe*
C.Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jan-2000
C.Accession: t8393
R.McLean, J.; Harris, D.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.
submitted to the EMBL Data Library, February 1996
A.Reference number: Z21791

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 21, 2002, 14:49:41 ; Search time 16.78 Seconds

(without alignments)
3188.940 Million cell updates/sec

Title: US-09-729-653-2

Perfect score: 7257
Sequence: 1 HSLIGRCSRLGDNVAVAC.....AAQLKRIIDCTOYLRFTEEL 1382

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	182.5	2.5	1653	1 CLH_YEAST	P22137 saccharomyc
2	177	2.4	1349	1 IK13_YEAST	O06706 saccharomyc
3	167	2.3	1332	1 IKAP_HUMAN	O95163 homo sapien
4	161	2.2	1666	1 CLH_SCHPO	Q10161 schizosacch
5	155	2.1	1455	1 DP2L_PSEAE	P42810 pseudomonas
6	155	2.1	1455	1 DP2L_PYRAB	O942f4 pyrococcus
7	152.5	2.1	2386	1 RAD3_SCHPO	O02099 schizosacch
8	151	2.1	4349	1 DYHC_FUSSO	P78716 fusarium so
9	149	2.1	561	1 SR72_SCHPO	O59787 schizosacch
10	148	2.0	1258	1 YS00_ANASP	O84702 anabaena sp
11	148	2.0	1557	1 DVAL_DICVI	O24702 dictyocaulu
12	147.5	2.0	1248	1 APAF_HUMAN	O14727 homo sapien
13	147	2.0	608	1 WDR1_XENLA	O947f2 xenopus lae
14	143.5	2.0	609	1 WDR1_CHICK	O93377 gallus gall
15	143.5	2.0	1681	1 CLH_CAEEL	P34574 caenorhabdi
16	142	2.0	608	1 WDR1_DROME	O94668 drosophila
17	141.5	1.9	1675	1 CLH_RAT	P11442 rattus norv
18	140.5	1.9	1675	1 CLH1_HUMAN	O00610 homo sapien
19	140.5	1.9	1675	1 CLH_BOVIN	P49951 bos taurus
20	140	1.9	2875	1 RRP1_TSWY1	P28976 tomato spot
21	138	1.9	567	1 YE28_METUA	O05823 methanococc
22	138	1.9	796	1 COPP_SCHPO	O42937 schizosacch
23	137.5	1.9	787	1 Y091_CAEEL	P41842 caenorhabdi
24	136.5	1.9	3056	1 ATP_HUMAN	Q13315 homo sapien
25	136	1.9	606	1 WDR1_HUMAN	O73083 homo sapien
26	136	1.9	1431	1 DP2L_PYRHO	O57861 pyrococcus
27	135	1.9	1249	1 APAF_RAT	O946v5 rattus norv
28	134.5	1.9	298	1 SNAB_BOVIN	P81126 bos taurus
29	134.5	1.9	298	1 SNAB_BOVIN	O94115 homo sapien
30	133.5	1.8	982	1 P115_MYCPY	O94115 mycoplasma
31	133	1.8	1657	1 T0G1_HUMAN	P46940 homo sapien
32	132.5	1.8	515	1 YCW2_YEAST	P25382 saccharomyc
33	132	1.8	606	1 WDR1_MOUSE	O88342 mus musculu

34	132	1.8	934	1 SYL1_SULSO	P58176 sulfolobus
35	131.5	1.8	1249	1 APAF_MOUSE	O88879 mus musculu
36	131	1.8	595	1 YB85_SCHPO	O14301 schizosacch
37	131	1.8	2547	1 PAFX_HUMAN	O93008 h probabile
38	130.5	1.8	1125	1 YB62_SCHPO	O14248 schizosacch
39	129.5	1.8	566	1 YG43_YEAST	P53197 saccharomyc
40	128	1.8	1150	1 YK01_CAEEL	O03560 caenorhabdi
41	127.5	1.8	897	1 DP01_HELPJ	O943e9 heliobacte
42	127	1.8	1263	1 DP2L_PYRPU	P81409 pyrococcus
43	127	1.8	1597	1 CTRO_MOUSE	P49025 mus musculu
44	126.5	1.7	716	1 CA80_EUPOC	O949x3 euploes oc
45	126.5	1.7	3354	1 CADN_MOUSE	O99pf4 mus musculu

ALIGNMENTS

```

RESULT 1
ID. CLH_YEAST STANDARD: PRT: 1653 AA.
AC P22137:
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Clathrin heavy chain.
GN CHC1 OR YGL206C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C;
RA MEDLINE=91093343; PubMed=1898742;
RA Lemmon S.K., Pelllicena-Palle A., Conley K., Freund C.L.;
RT "Sequence of the clathrin heavy chain from Saccharomyces cerevisiae
RT and requirement of the COOH terminus for clathrin function.";
RL J. Cell Biol. 112:65-80(1991).
RN [2]
RP SEQUENCE FROM N.A.
RA MEDLINE=97298309; PubMed=9153757;
RA Feuerhahn M., Simeonova L., Souciet J.-L., Potier S.;
RT "Analysis of 21.7 kb DNA sequence from the left arm of chromosome VII
RT reveals 11 open reading frames: two correspond to new genes.";
RL Yeast 13:475-477(1997).
CC - FUNCTION: CLATHRIN IS THE MAJOR PROTEIN OF THE POLYMERAL COAT OF
CC COATED PITS & VESICLES. IN YEAST, IT IS INVOLVED IN THE RETENTION
CC OF PROTEINS IN AN INTRACELLULAR MEMBRANE COMPARTMENT, PRESUMABLY
CC THE TRANS-GOLGI.
CC - SUBUNIT: CLATHRIN TRISKELIONS, COMPOSED OF 3 HEAVY CHAINS AND 3
CC LIGHT CHAINS, ARE THE BASIC SUBUNITS OF THE CLATHRIN COAT.
CC - SUBCELLULAR LOCATION: CYTOPLASMIC FACE OF COATED PITS AND
CC VESICLES.
CC - DOMAIN: THE C-TERMINAL THIRD OF THE HEAVY CHAINS FORMS THE HUB OF
CC THE TRISKELION. THIS REGION CONTAINS THE TRIMERIZATION DOMAIN AND
CC THE LIGHT-CHAIN BINDING DOMAIN INVOLVED IN THE ASSEMBLY OF THE
CC CLATHRIN LATTICE.
CC - SIMILARITY: BELONGS TO THE CLATHRIN HEAVY CHAIN FAMILY.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@sib-sib.ch).
-----
EMBL: X52900; CAA37082.1; -.
EMBL: 272728; CAA96919.1; -.
PIR: A36349; A36349.
HSP: P11442; IBBP.
SGD: S0003174; CHC1.
InterPro: IPR001473; Clathrin_propel.

```

DR	InterPro; IPRO00547; Clatrin_repeat.
DR	Pfam; PF01394; Clatrin_propel; 3.
DR	Pfam; PF00637; Clatrin_repeat; 7.
DR	SMART; SM00299; CLH; 7.
KW	Coated pits.
FT	DOMAIN 1 483 GLOBULAR TERMINAL DOMAIN.
FT	DOMAIN 484 527 FLEXIBLE LINKER.
FT	DOMAIN 528 1653 HEAVY CHAIN ARM.
FT	DOMAIN 528 ? DISTAL SEGMENT.
FT	DOMAIN ? 1653 PROXIMAL SEGMENT.
FT	DOMAIN 453 469 BINDING SITE FOR THE UNCOATING APPASE INVOLVED IN LATTICE DISASSEMBLY (POTENTIAL).
FT	BINDING 1219 1528 LIGHT CHAIN (BY SIMILARITY).
SQ	SEQUENCE 1653 AA; 187233 MW; 622F3083DFE2FE315 CRC64;

```

RA Yajima H Tokunaga M., Nakayama-Murayama A., Hisiluma F.:
RT "Characterization of IKIL and IK3 genes conferring PGKL killer
RL sensitivity on Saccharomyces cerevisiae."
RN Biosci. Biotechnol. Blochem. 61:704-709(1997).
RP (2)
RC SEQUENCE FROM N.A.
RA STRAIN-S288C / AB972;
RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Du Z.,
RA Favello A., Fulton L., Galtung S., Greco T., Kristen D., Kuaba T.,
RA Hallsworth K., Hawkins J., Hillier L., Jier M., Johnson D.,
RA Johnston L., Langston Y., Latreille P., Le T., Madis E., Menezes S.,
RA Miller N., Nhan M., Pauley A., Peluso D., Rifken L., Riles L.,
RA Tatch A., Trevasis E., Vignati D., Wilcox L., Wohlman P., Vaudin M.,
RA Wilson R., Waterston R.;
RL Submitted (Jan-1995) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: ESSENTIAL FOR THE INTRACELLULAR KILLING PROCESS OF PGKL
CC KILLER TOXIN.
CC -1- SUBCELLULAR LOCATION: MAY BE MEMBRANE ASSOCIATED.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D87841; BAA20120.1; -
DR EMBL; U19104; AAB67278.1; -
DR TRANSFAC; T03453; -
DR SGP; S0004376; IKI3.
KW Membrane.
SQ SEQUENCE 1349 AA; 152989 MW; 7AE81C0BA99282F6 CRC64;

Query Match 2.4%; Score 177; DB 1; Length 1349;
Best Local Similarity 18.2%; Pred. No. 0.0043;
Matches 245; Conservative 188; Mismatches 384; Indels 530; Gaps 66.

OY 10 GLLGD---GNVACVETGACVCLHSSRSSGSSVEKKRIFLSLEKTLGAPIOFAMOKTSG 66
DB 223 GLYGNQLDPTMPYAVDIDGDTYALDSHE-----ITISM-RGDC 259
OY 67 NYLAVTGYADYI-----VKIFDRHGQRSEIN-LPGNCVAMDWDKDGDLAV 111
DB 260 DYFAAVSVEEVDEDEDEFTSIKRAFRVFSREGQDLSSEPTVGMEHOLSMRPGCSLIAS 319
OY 112 IAEKSSCIYMDANFNKTSQLDNGARDQSFILMSKVGSLFVAGVYKGLXIIYNOTSK 171
DB 320 IQRKI-----DLGEDSDYVLIFFERNG-----LRHGEFDR 350
OY 172 IYVLAKHTKRIITCGCMNAENLXALGGEDEMTIYVNOEGDTIRGQVRSEPRXNMQFLMK 231
DB 351 LPL---DEKVSVCWNNS-----EALAVLVANR----- 376
OY 232 DDRTSAAESMTSVLGKTTFLPLNENPDNPADLEFQODFGNIVCY-NMYGGRIMIGFS 290
DB 377 -----IQLMYSKNYHWYK-----QELYASDISYVKHNRKDFLMS 414
OY 291 CGHFVYVISHGELQO-----EIFQARNKNDLITSLAV-SQTLNKVATCGDNCIKIODLVD 345
DB 415 DAGFINIVDFAYKMAOGPTLEP-----DNGTSLVVDGRIVNITP----- 454
OY 346 LKDMYVILNLDENKGLGLTSDDGQLLATSTGRSLHH-VFLKRLPIISGACS-TRIA 402
DB 455 -----LALANVPPPMYIRKPEFIPGNVLDAVACSNSNI 486
OY 403 YLTSLLEYTVANPV-----EGELPIYVSVDPERNFYAVGLYHLAVGMNRRAMPYVIGE 455
DB 487 YAIANKDVLIFAAVPSIEEMKKGHPISVCEPKESEFT----- 524
OY 456 NAYKKLKDMEYIG-TVASICLHSDYAAALFBEKVQVLIHLESETLDAQERRETRLPAYVD 514

```

Dd	525	SEVDSLQVQVAEINDSIYGVLLDD-----	NSRIALLDIQDITQPTLITIVYV	572
Qy	515	KCRILCHALNTSDF---LIYDTDTGVV-----	QYFYIEDW-QGVNDYRHHVSVKATFP	562
Dd	573	YDKIV--LIRSDFPYHNLVYETRDGFTVCQADAGQIMEITKFPQLVDRF----	YKR----	623
Qy	563	DPNGRLVFEIDKSDGVVYCPVMDATYELIPDSPTIKGYLVEMWPMDKGVFIATYDDKVV		622
Dd	624	-----	VHNTSAEDD-----	646
Qy	623	TYVEFKDTIGQAKVILAGST-----	KVPFAH-----	654
Dd	647	--TINNGKLRANQVLLASATYSLEITDSPLLEFTTAOHNIQFHNLSSTDKPLRYLVEEGVE		704
Qy	655	-----	NGELTCQOTSGKGVNNIYLTSGFLSNLKDGP----	701
Dd	705	DERVAIREGSGILVSVIPKSSVYLQATRG--	NLEFIYPRIMVLAEVR-----	757
Qy	702	FSDAMECR-----	LND-----	736
Dd	758	YKEAFVICTRHRLINDIHHYAPELFLENLEVINQIGRDYLNLFISCL--	SEDDVKT	815
Qy	737	YVRRAGNGIYMSLEQJKEGIEDNLLAGHLA-----	FTDYNLNAO--	786
Dd	816	KYKELTISGISKST---GHEPARLETMOYIKKKKKFDRKTSYVKNICDAVNLVLLSPE		871
Qy	787	-----	IAALERRDLQHMDSALQLAHLAPDQIPTISKEXAYI-OLEFAGD----	833
Dd	872	YKKKYQDITITAVASQNPQMSALAKIISELSE-----	EKOSCVYLCGFLQDVWVYKS	927
Qy	834	ALAHYEGCITGDNKEHDEACIAGYAQMSIR-----	MQDIRGVNQ	873
Dd	928	ALSTYDVSLA-----	LLVAAQSQMDPREYLPFLQELQDNEPLRKFLIDIDYLGANYEK	979
Qy	874	ALKHPSRYLKEDGCAILENNKQFSEAAQLEKGL--	YVD--	928
Dd	960	ALEHLSLEIDKP--GNVSEEVYIDVESHDLTKHGLALRYIDSEKQNVY--	NIYAK----	1033
Qy	929	LPHVSSPKRIHLQYAKAEADAGRYKEAVAVAYENAKQMSYIARYLIDLHLNPEKAVNIIVET		988
Dd	1031	--HLSSNQMYTDAAVAYEMIGKLKEAMGAVQSAKRWREMSIAVQGF--	PEVESVABEL	1086
Qy	969	QSLDGAKVVAFFLQLDQYSALQFLVMSKNNEATFTLAAQHNKMEITYADT--	IGSEDTTN	1047
Dd	1087	IS-----	SLTFEHRHYDA-ADIOLEVLDAVYK	1111
Qy	1048	EDYOSIALYFPEGKRYLQAGKFFLLCGQYSRALKHFLKPCSSSDNVAIMEALETVGOAKD		1107
Dd	1112	E-----	AVALY-----	1133
Qy	1108	ELTTNQLDHLGENDMCPKD--	AKYLFRLYMAIKOYREAAQTATIIAREEOSAGNYRNA	1165
Dd	1137	ELL--EEVVDPELGSGFGIINELLADCKGQINSQRLRLRE-----	LRAKKEENP-----	1183
Qy	1166	HDVLFSMYAEILKSQIKI--	PSEMAT	1189
Dd	1184	----	YAFYGQETEQADDVYASVAPSETSTP	1206
RESULT 3				
IKAP_HUMAN				
ID	IKAP_HUMAN	STANDARD:	PRT:	1332 AA.
AC	Q95163; Q9H327; Q9UG87;			
DT	16-OCT-2001 (Rel. 40, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	IkappaB kinase complex-associated protein (IKK complex-associated protein) (p150).			
GN	IKKAP OR IKAP.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catartini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			

[1] SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE-Cervical carcinoma;
RA MEDLINE=98421679; PubMed=9751059;
R Cohen L., Henzel W.J., Baenrele P.A.;
RT "IKAP is a scaffold protein of the IkappaB kinase complex.";
RL Nature 395:292-296(1998).
[2]
RN SEQUENCE FROM N.A., AND VARIANT FD PRO-696.
RP MEDLINE=21090531; PubMed=11179008;
RA Slungenhaupt S.A., Blumenfeld A., Gill S.P., Leyne M., Mull J.,
R Chajungco M.P., Liebert C.B., Chadwick B.P., Idelson M., Renik L.,
RA Robbins C.M., Makolowska I., Brownstein M.D., Krappmann D.,
RT "Tissue-specific expression of a splicing mutation in the IKKAP gene
causes familial dysautonomia.";
RL Am. J. Hum. Genet. 68:598-605(2001).
[3]
RN SEQUENCE OF 961-1332 FROM N.A.
RC TISSUE-Brain;
RA Mambutt R., Heubner D., Mewes H.-W., Gassenhuber J., Wiemann S.;
RL Submitted (May-1999) to the EMBL/GenBank/DBJ databases.
[4]
RN VARIANT FD PRO-696, AND EFFECT ON PHOSPHORYLATION.
RP MEDLINE=21090544; PubMed=11179021.
RA Anderson S.L., Coll R., Daly I.W., Kichula E.A., Rork M.J.,
R Volpi S.A., Ekstein J., Rubin B.Y.;
RT "Familial dysautonomia is caused by mutations of the IKAP gene.";
RL Am. J. Hum. Genet. 68:753-758(2001).
CC -1- FUNCTION: MAY ACT AS A SCAFFOLD PROTEIN THAT MAY ASSEMBLE ACTIVE
CC IKK-MAP3K14 COMPLEXES (IKKA, IKKB AND MAP3K14/NIK).
CC -1- SUBUNIT: INTERACTS PREFERENTIALLY WITH MAP3K14/NIK FOLLOWED BY
CC IKK-ALPHA AND IKK-BETA.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- PPM: PHOSPHORYLATED.
CC -1- DISEASE: DEFECTS IN IKKAP ARE THE CAUSE OF FAMILIAL DYSAUTONOMIA
CC (FD). ALSO KNOWN AS RILEY-DAY SYNDROME OR HEREDITARY SENSORY AND
CC AUTONOMIC NEUROPATHY TYPE III. THIS AUTOSOMAL RECESSIVE DISORDER
CC IS DUE TO THE POOR DEVELOPMENT AND SURVIVAL, AND PROGRESSIVE
CC DEGENERATION OF THE SENSORY, SYMPATHETIC AND PARASYMPATHETIC
CC NEURONS. FD INDIVIDUALS ARE AFFECTED WITH A VARIETY OF SYMPTOMS
CC SUCH AS DECREASED SENSITIVITY TO PAIN AND TEMPERATURE,
CC CARDIOVASCULAR INSTABILITY, RECURRENT PNEUMONIAS, VOMITING CRISES,
CC AND GASTROINTESTINAL DYSFUNCTION. IT IS PRIMARILY CONFINED TO
CC INDIVIDUALS OF ASHKENAZI JEWISH DESCENT, WITH AN INCIDENCE OF
CC 1/3600 LIVE BIRTHS.
CC -1- CAUTION: REF.3 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A
CC FRAMESHIFT IN POSITION 1286.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; AF044195; AAC64258.1; -;
DR EMBL; AF153419; AAG43369.1; -;
DR EMBL; AL049945; CAB43219.1; ALT_FRAME.
DR MIM; 603722; -;
DR MIM; 223900; -;
KW Phosphorylation; Disease mutation.
FT VARIANT 696 696 R -> P (IN FD; MILD PHENOTYPE;
FT PHOSPHORYLATION IS REDUCED).
FT /FTID=VAR_011327.
FT CONFLICT 304 304 R -> W (IN REF. 2).
FT CONFLICT 312 312 K -> E (IN REF. 2).
FT CONFLICT 754 754 P -> L (IN REF. 2).
FT CONFLICT 961 961 C -> G (IN REF. 3).
FT CONFLICT 1072 1072 S -> C (IN REF. 1).
FT CONFLICT 1158 1158 P -> L (IN REF. 3).
FT CONFLICT 1320 1320 I -> V (IN REF. 3).

SQ SEQUENCE 1332 AA; 150191 MW; 3FD65FAB554DA923 CRC64;
Query Match 2.3%; Score 167; DB 1; Length 1332;
Best Local Similarity 18.5%; Pred. No. 0.018;
Matches 250; Conservative 188; Mismatches 456; Indels 456; Gaps 65;
OY 51 TWLGAPIGFAMOKTSGNYLAV-----TGADYVRFDRHG--QKRSEINLPNGCVANDW 102
DB 199 FWRG-----DGFVAASVCPETGARK-VWVWNEFALQSTSE-PVAGLPALAW 246
OY 103 DKDGVLAIVAKSSCTILMDANTKKTQSLDNGMDQKSFILMSKVSFLAVGTAKNLX 162
DB 247 KPSGSLASTQDKP-----NQDDYFEERNG-----L 273
OY 163 IYNHOTSRIPIVGLKHTFRITCGCNAAE-NLXALGEDMKITVSNOEGTIRQTVRSEP 221
DB 274 LHGHT---LPEFL-KDEYKVNDLMMNASSVLALED-----LQREKSSITKTQVQWLT 324
OY 222 XNMOFFLKKMDRTSAESMISVILGKTLFPLNEDPNADLEFQDPFGNIVCYNNWG 281
DB 325 VGNHYMYLK-----QSLSFSTCGSKIVSL-KMDPVTYRLHVLCOGWHYLAIDWHW 375
OY 282 DGRIMIGSCGHFVYISHTGE-----LGEIFGARHKK 315
DB 376 TTRDSVGNSSDLSNVAVIDGNRLVTVPROTVPPMCTYOLLRPHPVNOYTFLAHQK 435
OY 316 DNLTSIAVSQTLNKVAT---CGDNCIKIDLV-----DLKDMVYLMD 356
DB 436 SN--DLAVLDASNQISYVKCGD-CPSADPTVGLAVGSGFVCLRTPLHERYKIQFEN 492
OY 357 EENKG---IGTISWTDGOLLALS---TQSGLHVFVTKPLPLDGCSTRIVYLSL 407
DB 493 NEDQVNLKLGTLTWIEEDVFLAVSHSEFSRPSYIH-HLT-----AASSEM----- 538
OY 408 LEVTYANPVEGLPTIVSDVPEPNVAVGLYHLAAGMNNRAFYVLENAVKKLDMETL 467
DB 539 -----DEBQGLNVSSAAND----- 554
OY 468 GTVASICLHS---DYAAALFEKGVLHLIESEILDQERETRLP----- 510
DB 555 GVITSLCCNSKTKSVVLQADQIKFLWESPFLAKIPKMNKSGRPVPRPCTQTELAM 614
OY 511 -----AVDKCR-----ILCHALTSPLIYGTGTGVQVFIYEDNQF----- 547
DB 615 IGEBCVGLGLDRCPREFINDIEVASNITSPFAVYDEFLLTTHSHRCQFCLEDAEFKTLQ 674
OY 548 ---VNDYRHPYSVKKIRPPDNGRLVFIDEKSDGFVYCPVNDATVYEIIPDFSPTIKGYLM 603
DB 675 AGLSSNNHVSHEGLVKV---ERGSRTIVV-VPODKVLVLMRGVLEVVHNAVLADT- 729
OY 604 ENMPMDKGVFLAYDDKVTYVYFHKDTIOGAKVILLASTKYPAFAHKPLLXNGELTCOTQ 663
DB 730 RKM-LDKLIMF-----KEAFECMRKLININPI-YDHNPKVFLGNVETFIKQ 773
OY 664 SKCVNNIYLTSHGFLSNLKDADPDELRLP---LAHNLMKRFSDAMEMCRILNDEAMN 719
DB 774 IDSVNHNIL---FETELKE---EDVTKTMYPAPVSSVYLSHDPGKNIDLVCDAMRAVM 827
OY 720 ELA---RACL---HHM-----EVEFAL-RVYRRTGN-----VGIYMSLEQIGIEDY 759
DB 828 ESINPHKVCISLTISHVAKKTPELEIVLOKVHELOGNAPSDPDVAVSABEALKYLLHVDV 887
OY 760 NLLACHLAFNTDYNLAODLYLASSCPAALERMRLDQHMDSALQLANHLADQIPIFSK 819
DB 888 NELYDH-SLGTGYDF---DLVL-----MVAEKSDKDPREYLPFLNTLKKMETNYORFTID 937
OY 820 EYALDLERAGDVYVNLAHYEKGTIDNK---EHDACLAGVAGMISIRMGDIRRGVNOALK 876
DB 938 KY-----LkYERKAIGHLSKCGPEYFPECNLTKDKNL-----YNEALK 976
OY 877 --HPSRVLRKDC---GATLENNKOFSEAAOLYERGLVYDKAASVYIRSKNNAKVGDLDP 930

```

Db 977 LVPSSQOYQDISIAYGEHLMQEHMEPEAGLMPARCAHEKALSAFLTCGMN----- 1028
Oy 931 HVSPKIHLYQAKAKADGRYKA--VVAENAKOMOSYIRYLDHINPEKAVNIREF 988
Db 1029 -----KQALCVAAQNLFTKDOLYGL-----GRTLAGKLVEDR 1060
Oy 989 OSLDGAKVNAFFLQDGYSAIOFLYKSCNNEAFTLAQHNKMEI----- 1035
Db 1061 KHIIDAAVLEE---SAQDYEEAVLLLEGAAWEEALRLVYKKNRDLIETNVPSTLEAQ 1117
Oy 1036 --YADITISEDIT-----NEDVQSIALLYEGEKERYIQAGKFFL----- 1071
Db 1118 KNVMAFFLDSQATFSPRHKRLRLVRELKEQAQOAGL--DDEVPHQGESDLFSETSSVVG 1175
Oy 1072 --LGGQYSA-----LKHFLKCPSSBDNVAIEMAIETVGO-----AKDE 1108
Db 1176 SMSKSKYHSNRSIRASRKKRKAERKKHSLKEGSPLEDLALLENLSVONTENLKE 1235
Oy 1109 LITNQLIDHLGENDMPKADAYLFRLYMALQYREAAQ---TAIIIR----- 1154
Db 1236 -----VYHIL-----KVLFLFEFDQGRELQAFEDTLQLMERSLPEITWLTLY 1278
Oy 1155 EEOAGCNTRNADVLFSMYAELEKSKIKIP 1184
Db 1279 QONSATPVLGPNSTANSIMASYOQKTSVP 1308

```

RESULT 4

CLH_SCHPO STANDARD: PRT: 1666 AA.

```

AC 010161;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Probable clathrin heavy chain.
GN SPAC26A3.05.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_Taxid=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA McLean J., Harris D., Bartell B.G., Rajandream M.A., Walsh S.V.;
RL Submitted (Feb-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: CLATHRIN IS THE MAJOR PROTEIN OF THE POLYHEDRAL COAT OF
CC COATED PITS AND VESICLES (BY SIMILARITY).
CC -1- SUBUNIT: CLATHRIN TRISKELIONS, COMPOSED OF 3 HEAVY CHAINS AND 3
CC LIGHT CHAINS, ARE THE BASIC SUBUNITS OF THE CLATHRIN COAT
CC (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC FACE OF COATED PITS AND
CC VESICLES (BY SIMILARITY).
CC -1- DOMAIN: THE C-TERMINAL THIRD OF THE HEAVY CHAINS FORMS THE HUB OF
CC THE TRISKELION. THIS REGION CONTAINS THE TRIMERIZATION DOMAIN AND
CC THE LIGHT-CHAIN BINDING DOMAIN INVOLVED IN THE ASSEMBLY OF THE
CC CLATHRIN LATTICE.
CC -1- SIMILARITY: BELONGS TO THE CLATHRIN HEAVY CHAIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: 269240; CAA93228.1; -.
DR HSSP: P11442; 1BPO.
DR InterPro: IPR001473; Clathrin_Propel.
DR InterPro: IPR000547; Clathrin_Repeat.
DR Pfam: PF01394; Clathrin_Propel.5.
DR Pfam: PF00637; Clathrin_repeat.7.

```

DR SMART: SM00299; CLH: 7.
 KW Hypothetical protein: Coated pits.
 SQ SEQUENCE 1666 AA; 190017 MW; 16A6E84FEA2A7EA9 CRC64;

Query Match 2.2%; Score 161; DB 1; Length 1666;
 Best Local Similarity 18.7%; Pred. No. 0.039;
 Matches 251; Conservative 167; Mismatches 445; Indels 476; Gaps 65;

```

Oy 192 LVALGEGDKMITVSNQEGTITQTVRSEPHXN-QEFLMKMDRISAASMS-----VVL 246
Db 15 LASVGIQSPSFGFANVTLESDDKYVCVCRDNPNGVNVQVYLDLDDPSVNLRRPISADSVILH 74
Oy 247 GKTLFELNLNBPNDPADLEFQDDFG-----NIVYNNYGGGRIMIGF-----SCGFVY 296
Db 75 PKKRIIALAQAQOLQVFDLEAKAKINSYMNQDVYWTWISDS--YIGAVTDTSVYHMTV 132
Oy 297 -----ISTHTGELGOEI--FOARNHKDNLTSIAVSQTLNKVATGDNCI-----KI- 340
Db 133 SGSDPVKMFDRHSLSNGTQIISYKSNYNEWFLLIGISSRDNRIA--GNLYSKRKYS 190
Oy 341 QDLVLKDMYVL--NLDEENKGLGTLSTDDQLLALSTQ--KSLHVFILKPLILG- 394
Db 191 QPLESHASAFAYIQEGVDHE-----VOYLALASRLPTGS-----KLSIVFY 232
Oy 395 DACSTRIAYLTSLEVTYVANPEGELPIYVSVDVEPNFAV-----GLYHL-----AVG 443
Db 233 DRNPNNPAPATITVDLFEPPPEAVNDPPIAIEIGSTYNAVAVYTKGFIHYVDETAKCIY 292
Oy 444 MN--NRAMFYVGENAVKRLKDMETLGTVASICLHSDYAAALFEGKVOLHLEISEILDA 500
Db 293 MNRVSGESIFVTYAHKSVNGLAINRKGVLSINPE-----TIIPYLSNLND- 342
Oy 501 QEERETRLPPAYDDCKRIICHALTSDFLYGTDTGVQVFIYEDMQFVNDYRIPSVYKTI 560
Db 343 -----PGL--AVRAMSHA-----NLPGADNLYMQGF--QQLAAGNYSEBAKVAAS 384
Oy 561 PDDPNGRTRLVFIDEKSDGVYVCVNDATYEIFDPSTIGVLNEMPMKGVFIAYDDK 620
Db 385 SP-----KGLRTSVDIQ----- 398
Oy 621 VYTYVFKHDTIQGAKYILAGSTKVPRAHKPLLYNSELTCQGTQSGKVNINYSTHGLSN 680
Db 399 -----FKLIQAAPGDI--APILQYFG-----T 418
Oy 681 LNDXGP-----DELRLAHNLMLKRSDAWEMKRIILDEAANWELARACLIHM----- 729
Db 419 LLDKGLNEHETIELARPVLAQ-----RIOLEKWKYGENKLACTEALGDLV 465
Oy 730 ---EVEFAIRVYRRIQNV--GIYMSLEQIKGIEDYMLGHLAMPNDYNLAQDYLASS 784
Db 466 KRYNPFALKIY-ETANVRKRYVWCLSE--LQDF-----GKLATYYSOONITPD----- 511
Oy 785 CPITALEMRRDLOHMSALQALAKHLAPDQIPFSKEAYALQLEFAGDVN----- 833
Db 512 -----YSLILQNLVRYNPDQ--AAEPATQMFMSNSINLEKLTVDIFMSQ 553
Oy 834 -----ALATYKGTIDGNKEHD-----EACLAGYAQMSIRNGDIRRGVONALKPSRY 881
Db 554 NLVQQAATLALDLAKDNDNEHSHLQTRLEINLINAPOVA---DAILG-NQMFTHFDRA 608
Oy 882 LKRDCAILENMKQFSEALQYKGL-YDDKASV-----YIRSKNNA----- 923
Db 609 V-----IASLCEKRGVQRALELYDKPADIRVIVHSNLNPEMLMTFSFSPD 658
Oy 924 KVGDLPLH-----VSSPKIHLYQAKAKADGRY----- 951
Db 659 EYVDYLREMLRSLNRLQNLQIVQIATRYSDLVQAQRIILEFKEFKFEGLYYVLGSIANI 718
Oy 952 ---KEAVVAEVA---KOMQSYIRIYLD-HLNNPKKAVNIYRETSIDGAKVNAFFLQ 1003
Db 719 TEDPEVYVRYIQAAACIMQNFTEVERICRDNNVNYNPEKVNKLKEARLADQLPLIL----- 773

```

```

QY 1004 LGDYSAIOFLVMSKCNNEAFLLAQQHNMETIA-----DITGSEDITTN 1047
Db 774 VCDRDVFNVDLVFLYFRNMFMQF-----IEIVQRINRSPKPPQVAGALLDIDCEELVQ 827
QY 1048 EDYOSIA-----LYFEGERK-----YLQAGKFFLLGQYSRAL-----1080
Db 828 NLMSVVGQVPVDELVEVEERNRRLKLLPYLES---LLQSSQDRALITDALAKIYIDSN 884
QY 1081 ---KHFLCPSSEEDNVAIEMAIETVQAKDELL-----TNOLIDLGENGMPKD 1128
Db 885 NMPEVFLK-----ENNFDYDPLTVGKYCKEKRDPYLAFLAYEKGNDPEIILNCNENSMFKQL 940
QY 1129 AKYLF-----LYMALKQ---YREAQTAIIITAREQSGNARNADVLFSTWAEKSK 1180
Db 941 AYLLKRLSDSNLMSSEVLDDSAZRPRLDQVIATVPESSDPAVSIVYALMEVDLPOL 1000
QY 1181 IKI-----PSEMATNLMLIHSYLLVKIHKNGDHMKGARMLIRANNISKFSPIVI 1233
Db 1001 IELLEKIYLOPSSFSSENNALQMLFLTAIKADS-----KMEYIDKLDKTYDVE 1050
QY 1234 LSTVIECHRAGLKNSAF-----SFAAMLREPEYRSKI 1266
Db 1051 IAEIATIE---NGLYEAEAFRIYKIHNNKEQAMKLVLEDIVSLDRAQDYAEVTPQPEVWSRL 1107
QY 1267 DAKYKKKIEGMWRPRDISF 1285
Db 1108 ---AKAQLDG-IRIPDAIE 1122

RESULT 5
YAG7_PSEAE STANDARD: PRT: 590 AA.
AC P42810; Q9HVC6;
DT 01-NOV-1995 (Rel. 32, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein PA4667.
GN PA4667.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PA01;
RX MEDLINE=95189718; PubMed=7883699;
RA Hungerer C., Troup B., Roemling U., Jahn D.;
RT "Regulation of the hema gene during 5-aminolevulinic acid formation
in Pseudomonas aeruginosa.";
RL J. Bacteriol. 177:1435-1443(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PA01;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltzy L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Iarbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Sailer M.H., Hancock R.E.W., Loy S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen.";
RL Nature 406:959-964(2000).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
PR EMBL; X82071; CAA57572.1; ALT_INIT.

```

```

DR EMBL; AE004880; AAG08054.1; -.
DR InterPro; IPR001440; TPR.
DR Pfam; PF00515; TPR; 5.
KW Hypothetical protein; Complete proteome.
FT CONFLICT 304 304 D -> DD (IN REF. 1).
SQ SEQUENCE 590 AA; 66290 MW; ABA9A48EB58497E CRC64;

Query Match 2.1%; Score 155; DB 1; Length 590;
Best Local Similarity 21.2%; Pred. No. 0.03;
Matches 133; Conservative 90; Mismatches 207; Indels 196; Gaps 34;

QY 759 YVLLAGHLAMFTNDYNLAODLYLASSCPYALAEMLRDLQHDLSAQLAKHL-----809
Db 69 YSLVLAELVGLQGNRRFDIALSNVVOA-----QKTRDPVSERRAFRIAYELDADDEALDT 122
QY 810 -----APDQIPFISKEVAIQLEFAGGVNVALAYEKGITGDNNKHEHDACLAGVAOMSI 862
Db 123 SLIMRSAPDNLD-AQRAAAYQLARAGRYEESWYMEKVLNQGQDTHFDLALSAETD- 180
QY 863 RMGDIRRGVNAQLKHPRYLKR--DGAILENMKQFSEAAQLYERK-----LYYDKAA 913
Db 181 --PDRFAGLQSFH---LLKYPNNQGL---FGKALLQDQGRDEALTLEDNSA 230
QY 914 S-----YIIRSK---NMAKVGDLP-----HYSSPKIHLYQAK-----AK- 945
Db 231 SHREVAFLRLRSRLQSMKRSDEALPLKAGIKRHPDKRVLAYARLVEQNRLDDKA 290
QY 946 -----EADGRYKAAVVAAYENAKOMQSVIRIYIDLHNPKEAVNIRETOSLGA 994
Db 291 EFAGLVQOFPDDDLRFSLALVCE-AQAKMDEA-RIVYEL-----VERDSHVDA 339
QY 995 KMAVRFLLQDGYSAIOFLVMSKCNNEAFLLAQQHNMKEIYADIGSEDITTN--EDYQS 1052
Db 340 -----HNLG-----RLAEE-----QKDTRALDEVAQ 362
QY 1053 IALYEGEKRYIQAGKFFLLCGYSRALKHFLKCPSSSEDNVAIEMAIETVQAKDELLTN 1112
Db 363 VGPNDFLPAQLQTDVLLKAGRVDEAAQRLDKASEQPDVAILQYL-----IEEAALSN 417
QY 1113 -----QLDHLIGENDGMPKDAKYLFRLYMALKQYREAAQ---TAIIAREQOSA- 1159
Db 418 NDQEKAMQALQEGKQ---YPEDNLLYTRSMLEKRNLDQMKDLRFYVAREPDNM 474
QY 1160 -----GNVRAADVLFESYALSKQIKIPSEMATNLMLIHSYLLVKIHKVN 1206
Db 475 ALNALGYTLADRTTRYGEARELI-----LKAHKLN-PDDPA-----ILDS-MGMINRQ 521
QY 1207 GDHMGARMLIRVANNISKFSPIHLYPILSTVIECHRAGLKNSAFSFAAMLREPEYRSKI 1266
Db 522 G-KLADARRYLQA--LQRYPDHEVAHLGEVYLAQ--GRQDA-----RAIWRFL 568
QY 1267 DAKYKKKIEGMWRPRDISIEEATTP 1292
Db 569 D-----KQPDSDVLRRTIKRLIGAEIP 590

RESULT 6
ID_DP2L_PYRAB STANDARD: PRT: 1455 AA.
AC Q9Y2F4;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE DNA polymerase II large subunit (EC 2.7.7.7) (Pol II) [contains: Pab
DE polC intein (Pab pol II intein)].
GN POLC OR PAB2404.
OS Pyrococcus abyssi.
OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.
OX NCBI_TaxID=29292;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ORSAY;

```

RA Heilig R.;
 RT "Pyrococcus abyssi genome sequence: Insights into archaeal chromosome
 structure and evolution";
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
 CC - FUNCTION: POSSESSES TWO ACTIVITIES: A DNA SYNTHESIS (POLYMERASE)
 CC AND AN EXONUCLEOTIC ACTIVITY THAT DEGRADATES SINGLE STRANDED DNA
 CC IN THE 3' TO 5' DIRECTION. HAS A TEMPLATE-PRIMER PREFERENCE WHICH
 CC IS CHARACTERISTIC OF A REPLICATIVE DNA POLYMERASE (BY SIMILARITY).
 CC - CATALYTIC ACTIVITY: N deoxynucleoside triphosphate - N diphosphate
 CC + (DNA)(N).
 CC - CATALYTIC ACTIVITY: DEGRADATION OF SINGLE-STRANDED DNA. IT ACTS
 CC PROGRESSIVELY IN A 3' TO 5' DIRECTION, RELEASING 5'-
 CC PHOSPHOMONONUCLEOTIDES.
 CC - SUBUNIT: HETERODIMER OF A LARGE SUBUNIT AND A SMALL SUBUNIT (BY
 CC SIMILARITY).
 CC - PTM: THIS PROTEIN UNDERGOES A PROTEIN SELF SPLICING THAT INVOLVES
 CC A POST-TRANSLATIONAL EXCISION OF THE INTERVENING REGION (INTEIN)
 CC FOLLOWED BY PEPTIDE LIGATION (POTENTIAL).
 CC - SIMILARITY: BELONGS TO THE ARCHAEAL DNA POLYMERASE II FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: AJ248283; CAB49044.1; -
 CC InterPro: IPR003586; HIntc.
 CC InterPro: IPR003587; HIntn.
 CC InterPro: IPR002203; Intein.
 CC SMART: SM00305; HIntc.1.
 CC SMART: SM00306; HIntn.1.
 CC PROSITE: PS00881; PROTEIN_SPLICING; FALSE_NEG.
 CC Transferrase: DNA-directed DNA polymerase; DNA replication; Hydrolase;
 CC Nuclease; Exonuclease; DNA-binding; Multifunctional enzyme;
 CC Autocatalytic cleavage; Protein splicing; Complete proteome.
 CC
 CC CHAIN 1 954 DNA POLYMERASE II LARGE SUBUNIT, 1ST PART
 CC (POTENTIAL)
 CC PAB POLC INTEIN (POTENTIAL).
 CC CHAIN 955 1139 DNA POLYMERASE II LARGE SUBUNIT, 2ND PART
 CC (POTENTIAL).
 CC CHAIN 1140 1455
 CC SEQUENCE 1455 AA; 165709 MW; BB092E7BB2FE0D21 CRC64;
 SO

Query Match 2.1%; Score 155; DB 1; Length 1455;
 Best Local Similarity 21.0%; Pred. No. 0.11;
 Matches 161; Conservative 90; Mismatches 288; Indels 226; Gaps 35;

QY 732 EFAIRRYRIGNVGYMSLEQIKGIEDYNL-----LAGHL 766
 DB 100 EQAVRTALALTEGIVSA--PIEGIANVKIKRTMTADNSEYALVYAGPIRSSGGTAQAL 157
 QY 767 AMFTNDY---NLADLYLASSCPIALEMKRDIOHMDSAQLAKHLAPDOIPTSKEVAT 823
 DB 158 SYLVGVYVRKRLGLDRFKPSKHEIEMVBEVDLYH-RAVTRLOHYHSPREVRAMRIPI 216
 QY 824 QLEFAGDYVNALAHYKGTGDKHEDEACLAGVAQMSIRMGDI---RCGVNALKHP 880
 DB 217 EI-----TGEATDDEVEVSHRDVPGVETNQLGAILVLAESVLQAKKLVK 262
 QY 881 VKRDCGAILENMKOPSEAAQLYE-----KGLYVKAASVYI 917
 DB 263 YIDKMGIEGHEWLTKEVEAKKEGPEEGKEESLAESTLEETVEVDMGITY---SLYQ 318
 QY 918 R-----SKMAK--VGDLPLPVSSPK---IHLQYAKAK-----945
 DB 319 KFEETIAPDSKVAKEVIGC-RPLESDPSKRGFRRLRGRASGAFATWGINTPATMLVDE 377
 QY 946 -----EADGRYKAENVATNAKQWQ-----SVIRIYLDHLNPKKAVNIYRET 988
 DB 378 FLAIGTOLKTERPG--KGAVVTPTTIEGPIVRLKDGSVLRV--DDYNLALKVREDVEEI 433

QY 989 QSLDGAKMVARFLOGLDVGSAIOLVMSKNNBAFTLAQOHKMEIYADIIGSEDTNE 1048
 DB 434 LYTGDA-----VIAGDFVENNOTLTPANYCEEMWLEEVYALKELTYE--VHLEPTEN 485
 QY 1049 DYOSIALYPERGERRYLOAGKFFLLCGQYSRALKHFLLKCPSSSDNVAIEMAI---ETVGOA 1105
 DB 486 EESI-----EASDYLEIDPEF-----LKEMLRDP-LRVKPPVELAIHFSEVLGIP 531
 QY 1106 KDELLTNQIDILLGENDCKPRDAKYLFRLYMAKQYREAAQTAAIIIAEEQSNAGNRA 1165
 DB 532 LHPYTT-----LYMNSVSPKDVEXLMRL--LKNYAEI-----EMSNF 566
 QY 1166 HDVLFMSVARELSOKIKIPSEMATNMLHSHYLLVYKHHVNG-----DHMGARMLIRVAN 1221
 DB 567 RGIKFAKVIYISOEKLGDGSKRTLELGLPHT-----VADGVYDYVPAALALPLPLGN 619
 QY 1222 NISKFPSHVPIILTSVYIECHRAGLKNARSF-AAMLMPEYRSKIDAK-----1269
 DB 620 LMNEFMAKPLIYATIDINENNEIKLRDGIISWIGARNGREKAKERKMKPPVGLPPIGL 679
 QY 1270 ---YKKIEGMVRRPDISLEIETATPCPFCKFLPPECCLCPGCKNSIPYCIATGRHMK 1326
 DB 680 AGCSSRDIKKAADEGKVAEVEIAFFKCPKCGHYGP--EHLCPNC-----GTRKELL- 728
 QY 1327 DDMTCVPHC--DFPALYSELKIMLNTESPSCMCSERLNAOQLKI 1369
 DB 729 --W-VCPRCNAEYPESOAE-----GIYVTCPCNVKALPYAKRKI 765

RESULT 7
 RAD3 SCHPO STANDARD; PRT; 2386 AA.
 ID RAD3 SCHPO 090UM1. Q92391;
 AC 01-JUL-1993 (Rel. 26, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE DNA repair protein rad3.
 GN RAD3 OR SPBC216.05.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomyces.
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
 RC STRAIN=972;
 RX MEDLINE=97133293; PubMed=8978690;
 RA Bentley N.J., Holtzman D.A., Flagg G., Keegan K.S., Demaggio A.,
 RA Ford J.C., Hoekstra M., Carr A.M.;
 RT "The Schizosaccharomyces pombe rad3 checkpoint gene.";
 RL EMBO J. 13:6641-6651(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RA Lyne M., Rajandream M.A., Barrell B.G., Oliver K., Harris D.;
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE OF 711-1781 FROM N.A.
 RX MEDLINE=93012978; PubMed=1398093;
 RA Seaton B.L., Yucel J., Sunnerhagen P., Subramani S.;
 RT "Isolation and characterization of the Schizosaccharomyces pombe rad3
 gene, involved in the DNA damage and DNA synthesis checkpoints.";
 RL gene.119:83-89(1992).
 CC - FUNCTION: INVOLVED IN G2 ARREST FOLLOWING DNA DAMAGE WHERE IT
 CC PHOSPHORYLATES CHK1. IT IS ALSO INVOLVED IN THE DEPENDENCE OF
 CC MITOSIS ON THE COMPLETION OF DNA REPLICATION.
 CC - SUBCELLULAR LOCATION: Nuclear.
 CC - SIMILARITY: BELONGS TO THE PI3/P14-KINASES FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: Y09076; CAA70297.1; -
 DR EMBL: U76307; AAC49607.1; -
 DR EMBL: AL049556; CAB40165.1; -
 DR EMBL: X63544; CAA45106.1; -
 DR PIR: S25834; S25834.
 DR InterPro: IPR003151; FATC.
 DR InterPro: IPR003152; FATC.
 DR InterPro: IPR000403; P13_P14_kinase.
 DR InterPro: IPR001440; TPR.
 DR Pfam: PF02259; FAT; 2.
 DR Pfam: PF02260; FATC; 1.
 DR Pfam: PF00454; P13_P14_kinase; 1.
 DR SMART: SM00146; P13K; 1.
 DR PROSITE: PS00915; P13_4_KINASE_1; FALSE_NEG.
 DR PROSITE: PS00916; P13_4_KINASE_2; FALSE_NEG.
 DR PROSITE: PS0290; P13_4_KINASE_3; 1.
 KW Transferrase; kinase; DNA damage; DNA synthesis; DNA repair;
 KW Nuclear protein.
 FT DOMAIN 2018 2326 P13K/P14K.
 FT CONFLICT 198 198 T -> S (IN REF. 1).
 FT CONFLICT 526 526 L -> P (IN REF. 1).
 FT CONFLICT 1779 1781 SEC -> LRM (IN REF. 3).
 SQ SEQUENCE 2386 AA; 273526 MW; DDEA343FFA3872D3 CRC64;

Query Match 2.1%; Score 152.5; DB 1; Length 2386;

Best Local Similarity 17.3%; Pred. No. 0.34; Mismatches 563; Indels 533; Gaps 78;

Matches 281; Conservative 244; Mismatches 563; Indels 533; Gaps 78;

QY 25 GCA-----CVLHSSSGSSVEMKRIFFSLKLT--WLGAPIO-----FAMQKTSQNY- 68
 DB 510 GCEVNSPFLDERSLFPIRYHELFCALKNPDISSSVKQSLLDGFRHSQHCSNN 569
 QY 69 -----LAVTG-----ADIVKIFDRHGQRSEINLPNCVAMDMDKDG-- 106
 DB 570 KESMLSLREFINKALASTSRCLRVAAAKVPIF-----IKGPNNDIYEFHESKA 620
 QY 107 ---DVLAVIAEKSSC-----IYLDANTNKTOLDNGMDOKSFLMSKVGSEFLAVGTAK 158
 DB 621 LIEFNLIKILAVENTAILLEVILLISLSRVVE-----EELHFVILEYSSVINSGIT 674
 QY 159 GNLXYIYNHQ---TSRRIPVLGKHTKRTTCGNAENALAGGEDKMTIVSN-----Q 207
 DB 675 QGIGLSALQOQIASTRHSIYW-----QLSPYWPVSAIVQGMGKKPNIASLFAQLMNS 729
 QY 208 EBDITIRQVNRSEPKNMOFLMKMDR---TSAESMISVVLGKTLPLFLNLE----- 258
 DB 730 EBDLIRIQVATLP-----FLVLTNKKALIVRIALSOSDV---ATLCLTNHKKILASLL 781
 QY 259 -PDNPADLEFQODFGNIVCYNMGDGRIMIGFSCGHF-----VYSTHTEGEOEFOCA 311
 DB 782 TTDPH-----NLESVMILLSTATSDPEKVDLTSLIRSPISITYELLQ 826
 QY 312 RNHKNLTSIAVSQTLNKVATGDCNCIKIQLDVLKDMYVILNDEENKGLTSLMTDGG 371
 DB 827 --YQNDVHEKIEINALRKVAMIVSQVNDDELSNKKELVDFN---NHILGILAEFSN- 879
 QY 372 QLLALSTORGLSHV-----FLTKLPILGDACSTRALVITLSLEVYAN----- 414
 DB 880 ---ILNDKGKTSINEKIKTIYIEKMLSLCGAVKLGIPQILSNISQAFQEHRLRFYAI 936
 QY 415 -----PVEGLPTIVSDVDPNPAVAGLVHL 440
 DB 937 KAWSLIILATKPEYSSTAGSLVLLPLFPLPEQAEALVQIR-----DTISSDTHNC 991
 QY 441 AVGNM-----NRAMFYVLGENAVKLLDMETLGTVASI--CLHSDYAAALFEKVOHL 491
 DB 992 LOGKMAIPTSLSDSACFSLKAKEIFCSLQNDYFSELSIICLNTNENPVCYGLQKLE 1051

QY 492 L-----ISEILDAQEEERTLFPVAVDKCRILCHALTSDFLIY----- 530
 DB 1052 LFFQAKVDELHPTLNDINSNEVD-----QLRLCLDCC--VKYASTNMQSYLAANK 1102
 QY 531 ---GT-----DTGVQVQFY--IEMQVYNDIRHVSVKKIFPPDNKIRLVE 571
 DB 1103 LSELGAIIDPSRAKAOHIIKETVYVLNDFENGESLSKILDFMQ---SOLIP---AFLVT 1154
 QY 572 IDEKSDGFVYCPVNDATYEIPDF-----SPTI---KG--VLMEWM-----PMDKCVFTA 615
 DB 1155 TDTKAGFL-----AYALOEFKLGKFSKAVINKKGLTYVTEHMSLPDLSKVLIP 1207
 QY 616 YDDDKVY-----TYVHKDTI-----QGA-----KVILA 639
 DB 1208 FLTSKYHLTPRIKIDIRYPIYKENVTIHTWMLFSLKLMETVAHSQNAEKIFGICSKYVD 1267
 QY 640 GSTKVPFAKRPILLVNGELTCTQTSQKVNNTYLTSHGFLSNLKDQRPDDLRLPLAHN--- 696
 DB 1268 QEVNIPCFLLPVLVNLVLT--ESELVNVK--IEEFQVLVNIQPPDGLNSVGQQRYS 1322
 QY 697 ---LMLKRFSDAWEMCRILNDEAAMNELARACLHMEVEFAIRYRRIGVNGIYSIE 751
 DB 1323 FVDVFEKIVDYLNKWLKMKKKNNMDSRAIAKENRYNVEDATSESSISKVESLSNF 1382
 QY 752 QIKGIEDVYLLAG-----HL-----AMFTNDYNLAQDLYLASSCP--IALE 791
 DB 1383 PSKTLGIYSLNCGFHARALFYWEQHIRNATAPAYAALESYRVLOETIYAGIDDPDEIEAVS 1442
 QY 792 M-----RRLQDQ-----WDSAL-----OLAKHLAPDQIFISKELAI-----QLEBAGDY 831
 DB 1443 LNFHDYSFDQQLLHNSGSTWDSALSCYEITIIQKDE-----NKKAKIILSNMLOSQHY 1497
 QY 832 VALAHYEKIGITGDNKEHDEACIAGV--AQMISIRMGDIRGVNOALKHSRYLKRDCGAI 889
 DB 1498 ESUVLSTDSFIIINDHREYKMLNLGIEASWRSLSIDSLKCLS--KSNLESFEAKLGI 1554
 QY 890 LENMKQFSAADLYEK--GLYDKAASVYIRSKNAKVG-----DLLPHVSEKITH 939
 DB 1555 FYQYLKDSFAELTERLOPLVDAATAI--ANGAHSAYDCYDIL-----SKLHA 1602
 QY 940 ---QYAKAKKAD-----GRYKEAV-----VAYENAKOMOSYIRIY 971
 DB 1603 INDFSRATIDTDIVSDNLDIVLRRLISOVAPRGKRRHQHLSHLVGYEKEFENTKTAELY 1662
 QY 972 LD-----HLNPD--EKAVINIVRETOSLDGAKV--ARFLOLDGYSALQIPLWMSKONE 1022
 DB 1663 LEIARISRKNGOFORAFNAILKAMDLDKPLATIEHQMMWMOHQHRAKATSELNFS-LNN 1721
 QY 1023 AFTLAQDQ-----NKEEIV-----ADIISEPTJNNDYOSIALYFEG 1059
 DB 1722 MFDLVDEHEERPKRKRETLGNPLKGVFLKTLKWLKQAQQLDKLDETYHHAVEIYSEC 1781
 QY 1060 ERRYLQAGKFFLLCGOYSALHFLKCPSESDNVAIEMALIEYVGOAKDILNTQILDHL 1119
 DB 1782 ENTHYLLGHRRLAMLEBEKLT-----FVNDQSEKFLSGELVTRTI 1821
 QY 1120 GE-----NDGMPKDAKYLFRLYMAL-----KQYREAQTAIIAR 1154
 DB 1822 NEFGRLSYLYGTNHIYESMK-----LTLMLDGAELRLSKDGEKYEFRH-----IIS 1872
 QY 1155 EEOASGNRYNADIVLFMYAEKLSOKIKIPSEMATNLMLTHSYILVKTIVKNG-----DH 1209
 DB 1873 RRRS-----LELMSNVCRLSMKIPQYFF--LVALSOMISRVCHPNKRYKILFH 1920
 QY 1210 MKGARMILIRVANNISKFPSPHIYPIILTSYV--ICHRAGAKNSAFSFAAMLRPREYSKIDA 1268
 DB 1921 I-----IANVASYEOETIMOLMATIKTSQKRSRLGSKSILVNLVHRSKLSMSKVDI 1972
 QY 1269 K 1269
 DB 1973 K 1973


```

RESULT      8
DHYC_FUSSO  STANDARD:      PRT: 4349 AA.
AC   p78716;
DT   01-NOV-1997 (rel. 35, Created)
DT   01-NOV-1997 (rel. 35, Last sequence update)
DT   30-NOV-2000 (rel. 39, Last annotation update)
DE   Dynein heavy chain, cytosolic (DHYC).
GN   DHCL.
OS   Fusarium solani (subsp. pist.) (Nectria haematococca).
OC   Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OX   Hypocreales; Nectriaceae; Haematonectria.
RN   NCBI_TaxID=109625;
RP   SEQUENCE FROM N.A.
RC   STRAIN=F213;
RA   Inoue S., Aist J.R., Turgeon B.G., Yoder O.C.;
RL   Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
CC   -1- FUNCTION: DYNEIN HAS ATPASE ACTIVITY. CYTOPLASMIC DYNEIN ACTS AS A
CC   MOTOR FOR THE INTRACELLULAR RETROGRADE MOTILITY OF VESICLES AND
CC   ORGANELLES ALONG MICROTUBULES.
CC   -1- SUBUNIT: CONSISTS OF AT LEAST TWO HEAVY CHAINS AND A NUMBER OF
CC   INTERMEDIATE AND LIGHT CHAINS.
CC   -1- SIMILARITY: BELONGS TO THE DYNEIN HEAVY CHAIN FAMILY.
CC   -----
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@sib-sib.ch).
CC   -----
DR   EMBL: U84215; AAC33176.1; -.
DR   HSSP: P03069; 1210.
DR   InterPro: IPR004273; Dynein_heavy.
DR   InterPro: IPR001482; GSPIL_E.
DR   InterPro: IPR002017; Spectrin.
DR   Pfam: PF03028; Dynein_heavy; 1.
KW   Motor protein; Microtubules; Dynein; ATP-binding; Coiled coil.
FT   DOMAIN 459      480      COILED COIL (POTENTIAL).
FT   DOMAIN 1178     1215     COILED COIL (POTENTIAL).
FT   DOMAIN 1266     1293     COILED COIL (POTENTIAL).
FT   DOMAIN 1334     1354     COILED COIL (POTENTIAL).
FT   DOMAIN 1560     1577     COILED COIL (POTENTIAL).
FT   DOMAIN 1640     1670     COILED COIL (POTENTIAL).
FT   DOMAIN 2048     2076     MICROTUBULE-BINDING (POTENTIAL).
FT   DOMAIN 2194     2217     COILED COIL (POTENTIAL).
FT   DOMAIN 3186     3284     COILED COIL (POTENTIAL).
FT   DOMAIN 3420     3477     COILED COIL (POTENTIAL).
FT   DOMAIN 3774     3807     COILED COIL (POTENTIAL).
FT   NP_BIND 1946     1953     ATP (POTENTIAL).
FT   NP_BIND 2239     2246     ATP (POTENTIAL).
FT   NP_BIND 2604     2611     ATP (POTENTIAL).
FT   NP_BIND 2946     2953     ATP (POTENTIAL).
SQ   SEQUENCE 4349 AA; 493453 MW; PCB3C7152B36A1BF CRC64;
Query Match      2.1%; Score 151; DB 1; Length 4349;
Best Local Similarity 18.2%; Pred. No. 1;
Matches 234; Conservative 172; Mismatches 432; Indels 446; Gaps 61;

```

```

QY 146 SKVGSFLAVGTGVKGNLXIYNHOTSRRKIPVIGKTHKRTTCGMAENMLALGDEKMTIVS 205
DB 613 SKRNALFVRPKIRGALQEOQONO-----LMHVKKO-----AINGLHEFKQO 653
QY 206 NOEGDTIRQTVRSEP--XNMOFFLKKMDRTSAESMISVYLKTKTFLFLNLNEPDNPA 263
DB 654 YGSHETNAQAQLDLDPVSGAIIWAKQIEFQLDGVRKKEAVLGPDMWHTTECHKLOERS 713
QY 264 DLEFQDFGNCIVCN-WYGD-GRIMFGSCGHHVIVSTHGTGELGQIFQARRNKKDNLTSI 321
DB 714 EL-EKOKLDPTARIYEAMIDVGRKISTIS-----GQ-LFE----- 746
QY 322 AVSQTNLKATCGDNCIKIQIDLVDLKDMYVILNLDENKGLTISWTDGQLLALSTORG 381
DB 747 -----IARVSAG-----GILELYNFPDPVYITLFEKTR 775
QY 382 SCHVFLTKLPLIGDACSTRAYLTSLLEVTVANPVGELPITYSVDEBNFAVGLYHLA 441
DB 776 NL-----TWOSYSVPHAVTVTSKDAKRVYPYAVSLMESVRTLSQTLRQIS 820
QY 442 VGNNNAMPYVIGENAVKLLKMEYIGTYASICLHSDYAALFEGVQOLH---LISEIL 498
DB 821 -----VMGEE-----SVLLFGYRNDVYKLLSEG--VPLRWESFINSHEL 857
QY 499 DAQEEETR-LFPAVDKCRILCHALTSDFLIYGTDTGVQVYVIDMQFVNDYRHPVSY 557
DB 858 FYSDNROTFRPLRG-----GTPGLAKNTESHKRGHPIRGEFAAVSY 898
QY 558 KIIEPDNGTRLVEIDEK--SDGEVYCPVNDATYEIIPDSPTIKGYLWENMPDKGVFA 615
DB 899 -----LOOKAVSLNFIHATVEQALKEL-----WTCPEEAAFNHS 932
QY 616 YDDDKYVTVFHKDTQGA--KYTLAGSTKVPRAHPRLLYNGLTCQOSQSGVNNIYLS 673
DB 933 -----RLDTTQAADQNLNLEQVYNLDF-----WAGLSKVOSILLTRLOSA 974
QY 674 THGFLSLNKDXGP-DELRPLAHN-----LMLKRFSD 704
DB 975 VHAMIEAFEDDTPDDDMRRKRVNNNNNEAKPDGPTMKRYVELAMRNOVIYLDLLEFAR 1034
QY 705 AWEM-----CRI-----LNDEAMNELARACLHMEVEFAIRYRR 740
DB 1035 SWFLHLHEWLGIVCNLRKIKATRYQMSLTFTTANDEPRFDLPSECAGLLQ-----RVY-- 1087
QY 741 IGVVGVMSLEQKIGIEDYNLAGHLAMPNDYNL--AQDLYLASSCPILAELKRPLOHW 799
DB 1088 ---VSVEKKLHEVSAYVDKWL-----OFQSLMDLQSEQVYDA-----LGEQLPRW 1129
QY 800 DSAIQAKHLAPQIPFISKEVAIQLEFAGDYVNALAH-----YEKGITGDNKEHD-----E 851
DB 1130 ---LQLLQETIKRSTFTDQVS-----RAGCHLTIDYDQVQTVYNAKYDQWQHE 1176
QY 852 ACLIAGVQMSIRMGDIRGVNQAQLKHPRYLKRDCGAILLENMKQFSEAAQVLEKGLYYDK 911
DB 1177 ILMKFSRSLGRNRRREINAEIEKARKH-----LES--QSSDASTAQAVQFITV 1222
QY 912 AASVYIRSKNMKAVGOLLPHVSSPKIHLQYAKAKKADGRKEAVVAYENAKQMSVIRYI 971
DB 1223 VQSCKRNVKTWAPETIDPFQGGSTLVROKQFPN--DWLHIEOI-----DSQWEAKETII 1275
QY 972 LDHLNPKAVNIVRETOSLDGAKNVARFFLQGDYSAIQFLVMSKCNNEAFT-LAQOH 1030
DB 1276 -----EKKSRIYQQTALQAKIYAE-----DKLINERIAELIAAQW 1311
QY 1031 NKMEIYADIIGSDTTNEDYQSTAL-YFEGEKRYLOAGKFFLLCGOYSRALKHLKCPSS 1089
DB 1312 NEKKPYVS-----GTQDPVASATLSFSRSRISKLODD-----AQMVAKAKKALDIPAS 1359
QY 1090 EDNVALEMAIEIV-----GQAKDELTT-----NQLIDHLEGDNG 1125
DB 1360 PD-TSLEATILEEVDQSVWSNLSITWASLNERRDLVMTAVQPRKIRSKYDIDLKSTKEK 1418
QY 1126 PKDAKYLFRLYMALKO-----YREAAQTALIIAREEOSAGNYRNAHDVLFPSMAELKSQRI 1181

```


FT REPEAT 1019 1059 MD 11.
 FT REPEAT 1060 1101 MD 12.
 FT REPEAT 1103 1143 MD 13.
 FT REPEAT 1144 1183 MD 14.
 FT REPEAT 1186 1227 MD 15.
 SQ SEQUENCE 1258 AA; 139513 MW; 45DF03B91170CA51 CRC64;

Query Match 2.0%; Score 148; DB 1; Length 1258;
 Best Local Similarity 19.3%; Pred. No. 0.25; Indels 178; Gaps 29;
 Matches 119; Conservative 89; Mismatches 232;

23 ETGCACVLHSSRGSSVEKRIFFSLLEKTLWLGAPIOFAMOKTSGNYLATGADYIVKIPD 82
 DB 661 DTDCHVRWVEVASKKLLICRHS-----NW-----VAFVFSFDGELLACGADENKWLMS 712
 QY 83 RHGQKRSEINLPNCV-----AMDWKDGDVLAIVAEKSSCIYLMADNTKTSQ 131
 DB 713 VR-----DGVCIKTLTGHEHEVEFSVAFHPDGETLA-SASGDKTKILMDIODGTCLQ 762
 QY 132 LQNGMRDQMSFILMSKVSFLAVGYKGNLXVYHOTSRIKIPVLCKHTKRITCGCMNAE- 190
 DB 763 TLTGHTDWRCVAFSPDQTLASSAADHTIKLMDVSQGCRLTKLSHTGWMNSVAFSADG 822
 QY 191 NLXALGGEDEKMTVSNQE-----GDTIRQTVRESEPMXNQFELMKMDPTSA---A 238
 DB 823 QTLASSGGRTIKIMNYHTGCECLKTYIGHTNSVYSAISP-DSKILVSGSGDRTIKLMDG 881
 QY 239 ESMISVILGKTKLFLINLNEPDNADLEFQDFGNIVCYN-----NY 280
 DB 882 QNHICL---KTLH---GHTNEVCSVAFSPDQTLACVSLDQSVRLMNCRTGQCLKAWY 933
 QY 281 G---DGRIMGFCGHRVYSTHTG-----ELGGEIFQARHKNLNSIAN---SQT 327
 DB 934 GNTDVALPVAFSPDQTLASSGNDKTVKLMWQTKYISSLEGHTDFIYGLAFSPDSQTL 993
 QY 328 NKVATCGDNCIKI-----QDLVLDKD-MYVIL----- 353
 DB 994 AASAT--DSVRLMNSTGOCFOILLEHTDWYAVVHFHQGIATGSDCYTKLKNIST 1051
 QY 354 ----NLDEENKGLGTLTMDGOLLALSTQSGSLHVFPLKLPILGDCASTRIAYL--- 404
 DB 1052 GQCLKTLSEHSDKILGMAMSPDQTLASSADQSVR-----LMDCTGRCVGLIRGH 1103
 QY 405 TSLLEVTVANPVEGELPIVSVDBEPNFAVVGILHLAVGMNNRMYVIGENAVKKLKMD 464
 DB 1104 SNRVSAISFSP-NGELIATCSTD-----QTVKIMWQOG---KCKLTL 1142
 QY 465 E-YLGTVASICLHSDYAALFEKGVOLHLESEILDQAEERETRLPVAVDKCRILC--- 520
 DB 1143 TGHNTNVPRDLAFSPD-----GKI-----LASASHDQIVRIMDVNTGKCHHICIGH 1187
 QY 521 -NALTSDFLIYGTGCV 537
 DB 1188 THLVSS--VAFSPDGEV 1203

RESULT 11
 DVAL_DICVI STANDARD; PRT; 1557 AA.
 AC 024702;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE DVA-1 polypeptide precursor (Antigen-1) (Allergen-1) (NPA).
 GN DVA-1.
 OS Dictyocaulus viviparus (Bovine Lungworm).
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;
 OC Trichostrongyloidea; Dictyocaulidae; Dictyocaulinae; Dictyocaulus.
 OX NCBI_TaxID:29172;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96089829; PubMed=8538702;

RA Britton C., Moore J., Gilleard J.S., Kennedy M.W.;
 RT "Extensive diversity in repeat unit sequences of the cDNA encoding
 RT the polypeptide antigen/allergen from the bovine lungworm
 RT Dictyocaulus viviparus.";
 RL Mol. Biochem. Parasitol. 72:77-88(1995).
 RN [2]
 RP FATTY ACID/RETINOID-BINDING.
 RX MEDLINE=95370256; PubMed=7642601;
 RA Kennedy M.W., Britton C., Price N.C., Kelly S.M., Cooper A.;
 RT "The DVA-1 polypeptide of the parasitic nematode Dictyocaulus
 RT viviparus. A small helix-rich lipid-binding protein.";
 RL J. Biol. Chem. 270:19277-19281(1995).
 CC -1- FUNCTION: HAS HIGH BINDING AFFINITY FOR FATTY ACIDS AND RETINOLIDS.
 CC -1- SIMILARITY: TO A.SUNM ABA-1.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 DR EMBL: U02568; AAC47403.1;
 KW Signal; Polypeptide; Lipid-binding; Retinol-binding; Repeat;
 KW Glycoprotein; Antigen; Allergen.
 FT CHAIN 1 21
 FT PEPTIDE 22 1557
 FT PEPTIDE 61 213
 FT PEPTIDE 214 334
 FT PEPTIDE 335 480
 FT PEPTIDE 481 606
 FT PEPTIDE 607 737
 FT PEPTIDE 738 869
 FT PEPTIDE 870 1003
 FT PEPTIDE 1004 1074
 FT PEPTIDE 1075 1208
 FT PEPTIDE 1209 1342
 FT PEPTIDE 1343 1475
 FT PEPTIDE 1476 1545
 FT PEPTIDE 1546 1557
 FT CARBOHYD 997 997
 FT VARIANT 946 946
 FT VARIANT 948 948
 FT VARIANT 954 955
 FT VARIANT 974 974
 FT VARIANT 990 990
 FT VARIANT 1543 1543
 SQ SEQUENCE 1557 AA; 182600 MW; 33717911CF55BDAB CRC64;

Query Match 2.0%; Score 148; DB 1; Length 1557;
 Best Local Similarity 18.6%; Pred. No. 0.35;
 Matches 256; Conservative 193; Mismatches 502; Indels 424; Gaps 71;

31 HSSRSQSSVEMKRIFFSLK--TWLGAPIOFAMOKTSGNYLATGGA---DIYKIPDRHG 85
 DB 408 HHHHRHLAVRRKRLVLAIEFLDMLKPE-----QKHLEKLENSGAHFDVIAEVKKFYG 462
 QY 86 ---QKRSEINLPNCVAMDMDK-----GDVLAIVAE----- 114
 DB 463 LPEEKIELKAKFKSCQCYDVAKEVATSEEMNDIMKHEKNSHDLKRLTELENRLTED 522
 QY 115 -----KSSCIYLMDA-NTNK--TSQDLNGMRDQMSFILMSKVSFLAVGYKGNLX 163
 DB 523 QKHTEIHRVCCGLMWQNTNOKHOSLEADAVLSMTDEDEKVKAIYETSNROT 562
 QY 164 YNH-----QTSRKIPVLGKHTKRITCGCMN-AENLXALGSD-KMITVSNQEGDTIRQTV 217
 DB 583 YDELTKMESSEDEVRKAKATEKLEAACKHGTNI--LGEENVDLIRMKKNGATFEISN 640
 QY 218 RSEPMXNQFELMK---MDRTSAESMISVILGKTKLFLINLNEPDNADLEFQO-DFGN 273

```

Db 641 RVDE-----LLEGITSDRREKAYRMSKLC--KRIYSLGSHK-----OLQOYDEEN 684
QY 274 IV--CYNW-----YDGR1-----MIGFCGHFVY 297
Db 685 VLQKLYTLWLDOSQKNELTMSDNKEIKYKIIDYFDGITGEVKEKAEVLOLACHNHYIK- 743
QY 298 STHGELCOEITFOARNHKNDLTSLA--VSQILNKVATGDCNCKIKOD----LVLDKDMVY 351
Db 744 SIVEGKMEIKQJLKECKSSSEIKAKVEDVYINQIS--DESIRSRADALLVCRIKFI 800
QY 352 ILNDEENKGGT-----LSWTDGOLLASTQSGSLHFLFKPLIGDAGCTRIAY 403
Db 801 VKRLRRDSSEIHSLEAMERITLWLSDOQYIKS-----IDVNDKRYLY 846
QY 404 LLSLEVTANPVEGELPTVSVDPENFVAVGLYALVGNMNRAMFVYLGEMAKKILMD 463
Db 847 -EKIME-----FDDAIGETK-----QKAKELKD 870
QY 464 MEYLGTVASICLHSDYAAALF---EGKVQHLIESEIL-DAQEEETRLFPAYVDKCR-- 517
Db 871 -----ACKH--YVVDLIGEENGNLIREKENGASNEALATKVEIMEIATIDETKRA 919
QY 518 ILCHALTSDPLITGDTGVVOYFYIEDMOFVNDYRHPVSVKRIEFDPRGTRLVFIDEKSD 577
Db 920 QAMASTSCRYV---GVVO-----RFRDHHENHDEALE---KHFTWLNEHOK 964
QY 578 GGVYCPVNDATYETIPDSPTIKGLVLENNMPDKGVFIAYDDKVYTYV-----FHKDT 630
Db 965 SOL-----KTIYESEDRBLHKV--WEFFEGAGLRSNASKKIYGVAKRFRDHHENH 1018
QY 631 IQGAKVILAGSTKVPFAKPLILYNGELTCOTQSGKVNNTIYSTGFLSNLKDQXPDELR 690
Db 1019 LDEA-----LEKYLTLWNEE---QKSOMKTY-----ESSDRE-- 1048
QY 691 PMLAHNLMLKRFSPAMKCRILNDEAMNELARACLIHM-----EVEFAIRVYRIGNV 744
Db 1049 -----ALYKVLFEFEATGEVKEKAVALKSACRHYIKDYGKREKKEIKEMESG-- 1100
QY 745 GIVMSLEOI-KGIEDYNILAGHLANFTNDYLAODLYLASSCPI--AALEMRDLQWMD 800
Db 1101 ---VSTEISIRKVDFF-----IAMITDEKKAKALRASSACKIYGVAKRFRDHHENH 1151
QY 801 SALQ-LAKHLAPDOIPETISKEVATOLEF---AGD-----YVALAHYKGTIGDNKEHDEA 852
Db 1152 NLEBPLEKTYLT-----WLNEQKSQMKTIYBSGDREALYKVLFEFF--AAEGVEKAAV 1205
QY 853 CIAGVAQMSIR--MGDIRGVNOALKHPSRYLKRDCGAILNEMKOFSEAAOLYEXGLYD 910
Db 1206 ELKSACRHYIDYIGDK-----AEKIKMKESGVSTEE----- 1239
QY 911 KAASVYIRSKMMAKVGDLPLHYSSPKHLOQAKAKEADGRYKEAVVAYENAKOMQSVIRI 970
Db 1240 -----ISK--KYDEFIAMITDE--KKAKALRASNAACK--IYVAKR---LRR 1279
QY 971 YLDHLNPEKRVNIVRETOISDGAKMVARFLOJGDYGSATQFLVMSKNNEAFTLAQO 1030
Db 1280 DHHHEHNLEEMAGXYLSWMSDQQAQVKKI-----YGTGR-----LATYN 1320
QY 1031 NKMEIYADIISEEDTJNEDYOSIALYFEGEKRYLOAGKFFLLCGOYRALKHFLKCPSE 1090
Db 1321 KVMLEFESVPDEK-----EK-----ATSLKKAACHRYITIDFQIK 1355
QY 1091 DNVAIEAIEFTVGOAKDELLTNOLIDHLLGENDGMPKDAKYLFLYMLAKOYREAAQTAI 1150
Db 1356 DNLAIVIKEMESG--ATNEAIGEKIDEFIAGLDDEOK-----AQOARASACKKI 1404
QY 1151 IIAEEDSAGRYRNAHVLFSMY-----AEKSKIKITPSEMATNLMLHSYILYK 1201
Db 1405 YGVASRRRREHYEIDVCAISKYLTLWNEEQAKELIKOLEK--DEKQF---IGKRIIEF 1458
QY 1202 IHVNGDHMKGARMILIRVANNISKFPSSHIVPLITSTVECHRAGLKNASFPAAM 1256
Db 1459 FELTSGDKEKARQOLKAA-----CKHYVKMYGEEKAALKILKIDSGISLEEM 1507

```

```

RESULT 12
ID APAF_HUMAN STANDARD. PRT: 1248 AA.
AC 014727; Q9UBZ5; Q9UB58; Q9UB59; Q9UB60; Q9UB61; Q9UB62; Q9UB63;
AC Q9UB64; Q9UB65; Q9UB66; Q9UB67; Q9UB68; Q9UB69; Q9UB70; Q9UB71;
AC Q9UB72; Q9UB73; Q9UB74; Q9UB75; Q9UB76; Q9UB77; Q9UB78; Q9UB79;
AC Q9UB80; Q9UB81; Q9UB82; Q9UB83; Q9UB84; Q9UB85; Q9UB86; Q9UB87;
AC Q9UB88; Q9UB89; Q9UB90; Q9UB91; Q9UB92; Q9UB93; Q9UB94; Q9UB95;
AC Q9UB96; Q9UB97; Q9UB98; Q9UB99; Q9UB100; Q9UB101; Q9UB102;
AC Q9UB103; Q9UB104; Q9UB105; Q9UB106; Q9UB107; Q9UB108; Q9UB109;
AC Q9UB110; Q9UB111; Q9UB112; Q9UB113; Q9UB114; Q9UB115; Q9UB116;
AC Q9UB117; Q9UB118; Q9UB119; Q9UB120; Q9UB121; Q9UB122; Q9UB123;
AC Q9UB124; Q9UB125; Q9UB126; Q9UB127; Q9UB128; Q9UB129; Q9UB130;
AC Q9UB131; Q9UB132; Q9UB133; Q9UB134; Q9UB135; Q9UB136; Q9UB137;
AC Q9UB138; Q9UB139; Q9UB140; Q9UB141; Q9UB142; Q9UB143; Q9UB144;
AC Q9UB145; Q9UB146; Q9UB147; Q9UB148; Q9UB149; Q9UB150; Q9UB151;
AC Q9UB152; Q9UB153; Q9UB154; Q9UB155; Q9UB156; Q9UB157; Q9UB158;
AC Q9UB159; Q9UB160; Q9UB161; Q9UB162; Q9UB163; Q9UB164; Q9UB165;
AC Q9UB166; Q9UB167; Q9UB168; Q9UB169; Q9UB170; Q9UB171; Q9UB172;
AC Q9UB173; Q9UB174; Q9UB175; Q9UB176; Q9UB177; Q9UB178; Q9UB179;
AC Q9UB180; Q9UB181; Q9UB182; Q9UB183; Q9UB184; Q9UB185; Q9UB186;
AC Q9UB187; Q9UB188; Q9UB189; Q9UB190; Q9UB191; Q9UB192; Q9UB193;
AC Q9UB194; Q9UB195; Q9UB196; Q9UB197; Q9UB198; Q9UB199; Q9UB200;
AC Q9UB201; Q9UB202; Q9UB203; Q9UB204; Q9UB205; Q9UB206; Q9UB207;
AC Q9UB208; Q9UB209; Q9UB210; Q9UB211; Q9UB212; Q9UB213; Q9UB214;
AC Q9UB215; Q9UB216; Q9UB217; Q9UB218; Q9UB219; Q9UB220; Q9UB221;
AC Q9UB222; Q9UB223; Q9UB224; Q9UB225; Q9UB226; Q9UB227; Q9UB228;
AC Q9UB229; Q9UB230; Q9UB231; Q9UB232; Q9UB233; Q9UB234; Q9UB235;
AC Q9UB236; Q9UB237; Q9UB238; Q9UB239; Q9UB240; Q9UB241; Q9UB242;
AC Q9UB243; Q9UB244; Q9UB245; Q9UB246; Q9UB247; Q9UB248; Q9UB249;
AC Q9UB250; Q9UB251; Q9UB252; Q9UB253; Q9UB254; Q9UB255; Q9UB256;
AC Q9UB257; Q9UB258; Q9UB259; Q9UB260; Q9UB261; Q9UB262; Q9UB263;
AC Q9UB264; Q9UB265; Q9UB266; Q9UB267; Q9UB268; Q9UB269; Q9UB270;
AC Q9UB271; Q9UB272; Q9UB273; Q9UB274; Q9UB275; Q9UB276; Q9UB277;
AC Q9UB278; Q9UB279; Q9UB280; Q9UB281; Q9UB282; Q9UB283; Q9UB284;
AC Q9UB285; Q9UB286; Q9UB287; Q9UB288; Q9UB289; Q9UB290; Q9UB291;
AC Q9UB292; Q9UB293; Q9UB294; Q9UB295; Q9UB296; Q9UB297; Q9UB298;
AC Q9UB299; Q9UB300; Q9UB301; Q9UB302; Q9UB303; Q9UB304; Q9UB305;
AC Q9UB306; Q9UB307; Q9UB308; Q9UB309; Q9UB310; Q9UB311; Q9UB312;
AC Q9UB313; Q9UB314; Q9UB315; Q9UB316; Q9UB317; Q9UB318; Q9UB319;
AC Q9UB320; Q9UB321; Q9UB322; Q9UB323; Q9UB324; Q9UB325; Q9UB326;
AC Q9UB327; Q9UB328; Q9UB329; Q9UB330; Q9UB331; Q9UB332; Q9UB333;
AC Q9UB334; Q9UB335; Q9UB336; Q9UB337; Q9UB338; Q9UB339; Q9UB340;
AC Q9UB341; Q9UB342; Q9UB343; Q9UB344; Q9UB345; Q9UB346; Q9UB347;
AC Q9UB348; Q9UB349; Q9UB350; Q9UB351; Q9UB352; Q9UB353; Q9UB354;
AC Q9UB355; Q9UB356; Q9UB357; Q9UB358; Q9UB359; Q9UB360; Q9UB361;
AC Q9UB362; Q9UB363; Q9UB364; Q9UB365; Q9UB366; Q9UB367; Q9UB368;
AC Q9UB369; Q9UB370; Q9UB371; Q9UB372; Q9UB373; Q9UB374; Q9UB375;
AC Q9UB376; Q9UB377; Q9UB378; Q9UB379; Q9UB380; Q9UB381; Q9UB382;
AC Q9UB383; Q9UB384; Q9UB385; Q9UB386; Q9UB387; Q9UB388; Q9UB389;
AC Q9UB390; Q9UB391; Q9UB392; Q9UB393; Q9UB394; Q9UB395; Q9UB396;
AC Q9UB397; Q9UB398; Q9UB399; Q9UB400; Q9UB401; Q9UB402; Q9UB403;
AC Q9UB404; Q9UB405; Q9UB406; Q9UB407; Q9UB408; Q9UB409; Q9UB410;
AC Q9UB411; Q9UB412; Q9UB413; Q9UB414; Q9UB415; Q9UB416; Q9UB417;
AC Q9UB418; Q9UB419; Q9UB420; Q9UB421; Q9UB422; Q9UB423; Q9UB424;
AC Q9UB425; Q9UB426; Q9UB427; Q9UB428; Q9UB429; Q9UB430; Q9UB431;
AC Q9UB432; Q9UB433; Q9UB434; Q9UB435; Q9UB436; Q9UB437; Q9UB438;
AC Q9UB439; Q9UB440; Q9UB441; Q9UB442; Q9UB443; Q9UB444; Q9UB445;
AC Q9UB446; Q9UB447; Q9UB448; Q9UB449; Q9UB450; Q9UB451; Q9UB452;
AC Q9UB453; Q9UB454; Q9UB455; Q9UB456; Q9UB457; Q9UB458; Q9UB459;
AC Q9UB460; Q9UB461; Q9UB462; Q9UB463; Q9UB464; Q9UB465; Q9UB466;
AC Q9UB467; Q9UB468; Q9UB469; Q9UB470; Q9UB471; Q9UB472; Q9UB473;
AC Q9UB474; Q9UB475; Q9UB476; Q9UB477; Q9UB478; Q9UB479; Q9UB480;
AC Q9UB481; Q9UB482; Q9UB483; Q9UB484; Q9UB485; Q9UB486; Q9UB487;
AC Q9UB488; Q9UB489; Q9UB490; Q9UB491; Q9UB492; Q9UB493; Q9UB494;
AC Q9UB495; Q9UB496; Q9UB497; Q9UB498; Q9UB499; Q9UB500; Q9UB501;
AC Q9UB502; Q9UB503; Q9UB504; Q9UB505; Q9UB506; Q9UB507; Q9UB508;
AC Q9UB509; Q9UB510; Q9UB511; Q9UB512; Q9UB513; Q9UB514; Q9UB515;
AC Q9UB516; Q9UB517; Q9UB518; Q9UB519; Q9UB520; Q9UB521; Q9UB522;
AC Q9UB523; Q9UB524; Q9UB525; Q9UB526; Q9UB527; Q9UB528; Q9UB529;
AC Q9UB530; Q9UB531; Q9UB532; Q9UB533; Q9UB534; Q9UB535; Q9UB536;
AC Q9UB537; Q9UB538; Q9UB539; Q9UB540; Q9UB541; Q9UB542; Q9UB543;
AC Q9UB544; Q9UB545; Q9UB546; Q9UB547; Q9UB548; Q9UB549; Q9UB550;
AC Q9UB551; Q9UB552; Q9UB553; Q9UB554; Q9UB555; Q9UB556; Q9UB557;
AC Q9UB558; Q9UB559; Q9UB560; Q9UB561; Q9UB562; Q9UB563; Q9UB564;
AC Q9UB565; Q9UB566; Q9UB567; Q9UB568; Q9UB569; Q9UB570; Q9UB571;
AC Q9UB572; Q9UB573; Q9UB574; Q9UB575; Q9UB576; Q9UB577; Q9UB578;
AC Q9UB579; Q9UB580; Q9UB581; Q9UB582; Q9UB583; Q9UB584; Q9UB585;
AC Q9UB586; Q9UB587; Q9UB588; Q9UB589; Q9UB590; Q9UB591; Q9UB592;
AC Q9UB593; Q9UB594; Q9UB595; Q9UB596; Q9UB597; Q9UB598; Q9UB599;
AC Q9UB600; Q9UB601; Q9UB602; Q9UB603; Q9UB604; Q9UB605; Q9UB606;
AC Q9UB607; Q9UB608; Q9UB609; Q9UB610; Q9UB611; Q9UB612; Q9UB613;
AC Q9UB614; Q9UB615; Q9UB616; Q9UB617; Q9UB618; Q9UB619; Q9UB620;
AC Q9UB621; Q9UB622; Q9UB623; Q9UB624; Q9UB625; Q9UB626; Q9UB627;
AC Q9UB628; Q9UB629; Q9UB630; Q9UB631; Q9UB632; Q9UB633; Q9UB634;
AC Q9UB635; Q9UB636; Q9UB637; Q9UB638; Q9UB639; Q9UB640; Q9UB641;
AC Q9UB642; Q9UB643; Q9UB644; Q9UB645; Q9UB646; Q9UB647; Q9UB648;
AC Q9UB649; Q9UB650; Q9UB651; Q9UB652; Q9UB653; Q9UB654; Q9UB655;
AC Q9UB656; Q9UB657; Q9UB658; Q9UB659; Q9UB660; Q9UB661; Q9UB662;
AC Q9UB663; Q9UB664; Q9UB665; Q9UB666; Q9UB667; Q9UB668; Q9UB669;
AC Q9UB670; Q9UB671; Q9UB672; Q9UB673; Q9UB674; Q9UB675; Q9UB676;
AC Q9UB677; Q9UB678; Q9UB679; Q9UB680; Q9UB681; Q9UB682; Q9UB683;
AC Q9UB684; Q9UB685; Q9UB686; Q9UB687; Q9UB688; Q9UB689; Q9UB690;
AC Q9UB691; Q9UB692; Q9UB693; Q9UB694; Q9UB695; Q9UB696; Q9UB697;
AC Q9UB698; Q9UB699; Q9UB700; Q9UB701; Q9UB702; Q9UB703; Q9UB704;
AC Q9UB705; Q9UB706; Q9UB707; Q9UB708; Q9UB709; Q9UB710; Q9UB711;
AC Q9UB712; Q9UB713; Q9UB714; Q9UB715; Q9UB716; Q9UB717; Q9UB718;
AC Q9UB719; Q9UB720; Q9UB721; Q9UB722; Q9UB723; Q9UB724; Q9UB725;
AC Q9UB726; Q9UB727; Q9UB728; Q9UB729; Q9UB730; Q9UB731; Q9UB732;
AC Q9UB733; Q9UB734; Q9UB735; Q9UB736; Q9UB737; Q9UB738; Q9UB739;
AC Q9UB740; Q9UB741; Q9UB742; Q9UB743; Q9UB744; Q9UB745; Q9UB746;
AC Q9UB747; Q9UB748; Q9UB749; Q9UB750; Q9UB751; Q9UB752; Q9UB753;
AC Q9UB754; Q9UB755; Q9UB756; Q9UB757; Q9UB758; Q9UB759; Q9UB760;
AC Q9UB761; Q9UB762; Q9UB763; Q9UB764; Q9UB765; Q9UB766; Q9UB767;
AC Q9UB768; Q9UB769; Q9UB770; Q9UB771; Q9UB772; Q9UB773; Q9UB774;
AC Q9UB775; Q9UB776; Q9UB777; Q9UB778; Q9UB779; Q9UB780; Q9UB781;
AC Q9UB782; Q9UB783; Q9UB784; Q9UB785; Q9UB786; Q9UB787; Q9UB788;
AC Q9UB789; Q9UB790; Q9UB791; Q9UB792; Q9UB793; Q9UB794; Q9UB795;
AC Q9UB796; Q9UB797; Q9UB798; Q9UB799; Q9UB800; Q9UB801; Q9UB802;
AC Q9UB803; Q9UB804; Q9UB805; Q9UB806; Q9UB807; Q9UB808; Q9UB809;
AC Q9UB810; Q9UB811; Q9UB812; Q9UB813; Q9UB814; Q9UB815; Q9UB816;
AC Q9UB817; Q9UB818; Q9UB819; Q9UB820; Q9UB821; Q9UB822; Q9UB823;
AC Q9UB824; Q9UB825; Q9UB826; Q9UB827; Q9UB828; Q9UB829; Q9UB830;
AC Q9UB831; Q9UB832; Q9UB833; Q9UB834; Q9UB835; Q9UB836; Q9UB837;
AC Q9UB838; Q9UB839; Q9UB840; Q9UB841; Q9UB842; Q9UB843; Q9UB844;
AC Q9UB845; Q9UB846; Q9UB847; Q9UB848; Q9UB849; Q9UB850; Q9UB851;
AC Q9UB852; Q9UB853; Q9UB854; Q9UB855; Q9UB856; Q9UB857; Q9UB858;
AC Q9UB859; Q9UB860; Q9UB861; Q9UB862; Q9UB863; Q9UB864; Q9UB865;
AC Q9UB866; Q9UB867; Q9UB868; Q9UB869; Q9UB870; Q9UB871; Q9UB872;
AC Q9UB873; Q9UB874; Q9UB875; Q9UB876; Q9UB877; Q9UB878; Q9UB879;
AC Q9UB880; Q9UB881; Q9UB882; Q9UB883; Q9UB884; Q9UB885; Q9UB886;
AC Q9UB887; Q9UB888; Q9UB889; Q9UB890; Q9UB891; Q9UB892; Q9UB893;
AC Q9UB894; Q9UB895; Q9UB896; Q9UB897; Q9UB898; Q9UB899; Q9UB900;
AC Q9UB901; Q9UB902; Q9UB903; Q9UB904; Q9UB905; Q9UB906; Q9UB907;
AC Q9UB908; Q9UB909; Q9UB910; Q9UB911; Q9UB912; Q9UB913; Q9UB914;
AC Q9UB915; Q9UB916; Q9UB917; Q9UB918; Q9UB919; Q9UB920; Q9UB921;
AC Q9UB922; Q9UB923; Q9UB924; Q9UB925; Q9UB926; Q9UB927; Q9UB928;
AC Q9UB929; Q9UB930; Q9UB931; Q9UB932; Q9UB933; Q9UB934; Q9UB935;
AC Q9UB936; Q9UB937; Q9UB938; Q9UB939; Q9UB940; Q9UB941; Q9UB942;
AC Q9UB943; Q9UB944; Q9UB945; Q9UB946; Q9UB947; Q9UB948; Q9UB949;
AC Q9UB950; Q9UB951; Q9UB952; Q9UB953; Q9UB954; Q9UB955; Q9UB956;
AC Q9UB957; Q9UB958; Q9UB959; Q9UB960; Q9UB961; Q9UB962; Q9UB963;
AC Q9UB964; Q9UB965; Q9UB966; Q9UB967; Q9UB968; Q9UB969; Q9UB970;
AC Q9UB971; Q9UB972; Q9UB973; Q9UB974; Q9UB975; Q9UB976; Q9UB977;
AC Q9UB978; Q9UB979; Q9UB980; Q9UB981; Q9UB982; Q9UB983; Q9UB984;
AC Q9UB985; Q9UB986; Q9UB987; Q9UB988; Q9UB989; Q9UB990; Q9UB991;
AC Q9UB992; Q9UB993; Q9UB994; Q9UB995; Q9UB996; Q9UB997; Q9UB998;
AC Q9UB999; Q9UB1000; Q9UB1001; Q9UB1002; Q9UB1003; Q9UB1004;
AC Q9UB1005; Q9UB1006; Q9UB1007; Q9UB1008; Q9UB1009; Q9UB1010;
AC Q9UB1011; Q9UB1012; Q9UB1013; Q9UB1014; Q9UB1015; Q9UB1016;
AC Q9UB1017; Q9UB1018; Q9UB1019; Q9UB1020; Q9UB1021; Q9UB1022;
AC Q9UB1023; Q9UB1024; Q9UB1025; Q9UB1026; Q9UB1027; Q9UB1028;
AC Q9UB1029; Q9UB1030; Q9UB1031; Q9UB1032; Q9UB1033; Q9UB1034;
AC Q9UB1035; Q9UB1036; Q9UB1037; Q9UB1038; Q9UB1039; Q9UB1040;
AC Q9UB1041; Q9UB1042; Q9UB1043; Q9UB1044; Q9UB1045; Q9UB1046;
AC Q9UB1047; Q9UB1048; Q9UB1049; Q9UB1050; Q9UB1051; Q9UB1052;
AC Q9UB1053; Q9UB1054; Q9UB1055; Q9UB1056; Q9UB1057; Q9UB1058;
AC Q9UB1059; Q9UB1060; Q9UB1061; Q9UB1062; Q9UB1063; Q9UB1064;
AC Q9UB1065; Q9UB1066; Q9UB1067; Q9UB1068; Q9UB1069; Q9UB1070;
AC Q9UB1071; Q9UB1072; Q9UB1073; Q9UB1074; Q9UB1075; Q9UB1076;
AC Q9UB1077; Q9UB1078; Q9UB1079; Q9UB1080; Q9UB1081; Q9UB1082;
AC Q9UB1083; Q9UB1084; Q9UB1085; Q9UB1086; Q9UB1087; Q9UB1088;
AC Q9UB1089; Q9UB1090; Q9UB1091; Q9UB1092; Q9UB1093; Q9UB1094;
AC Q9UB1095; Q9UB1096; Q9UB1097; Q9UB1098; Q9UB1099; Q9UB1100;
AC Q9UB1101; Q9UB1102; Q9UB1103; Q9UB1104; Q9UB1105; Q9UB1106;
AC Q9UB1107; Q9UB1108; Q9UB1109; Q9UB1110; Q9UB1111; Q9UB1112;
AC Q9UB1113; Q9UB1114; Q9UB1115; Q9UB1116; Q9UB1117; Q9UB1118;
AC Q9UB1119; Q9UB1120; Q9UB1121; Q9UB1122; Q9UB1123; Q9UB1124;
AC Q9UB1125; Q9UB1126; Q9UB1127; Q9UB1128; Q9UB1129; Q9UB1130;
AC Q9UB1131; Q9UB1132; Q9UB1133; Q9UB1134; Q9UB1135; Q9UB1136;
AC Q9UB1137; Q9UB1138; Q9UB1139; Q9UB1140; Q9UB1141; Q9UB1142;
AC Q9UB1143; Q9UB1144; Q9UB1145; Q9UB1146; Q9UB1147; Q9UB1148;
AC Q9UB1149; Q9UB1150; Q9UB1151; Q9UB1152; Q9UB1153; Q9UB1154;
AC Q9UB1155; Q9UB1156; Q9UB1157; Q9UB1158; Q9UB1159; Q9UB1160;
AC Q9UB1161; Q9UB1162; Q9UB1163; Q9UB1164; Q9UB1165; Q9UB1166;
AC Q9UB1167; Q9UB1168; Q9UB1169; Q9UB1170; Q9UB1171; Q9UB1172;
AC Q9UB1173; Q9UB1174; Q9UB1175; Q9UB1176; Q9UB1177; Q9UB1178;
AC Q9UB1179; Q9UB1180; Q9UB1181; Q9UB1182; Q9UB1183; Q9UB1184;
AC Q9UB1185; Q9UB1186; Q9UB1187; Q9UB1188; Q9UB1189; Q9UB1190;
AC Q9UB1191; Q9UB1192; Q9UB1193; Q9UB1194; Q9UB1195; Q9UB1196;
AC Q9UB1197; Q9UB1198; Q9UB1199; Q9UB1200; Q9UB1201; Q9UB1202;
AC Q9UB1203; Q9UB1204; Q9UB1205; Q9UB1206; Q9UB1207; Q9UB1208;
AC Q9UB1209; Q9UB1210; Q9UB1211; Q9UB1212; Q9UB1213; Q9UB1214;
AC Q9UB1215; Q9UB1216; Q9UB1217; Q9UB1218; Q9UB1219; Q9UB1220;
AC Q9UB1221; Q9UB1222; Q9UB1223; Q9UB1224; Q9UB1225; Q9UB1226;
AC Q9UB1227; Q9UB1228; Q9UB1229; Q9UB1230; Q9UB1231; Q9UB1232;
AC Q9UB1233; Q9UB1234; Q9UB1235; Q9UB1236; Q9UB1237; Q9UB1238;
AC Q9UB1239; Q9UB1240; Q9UB1241; Q9UB1242; Q9UB1243; Q9UB1244;
AC Q9UB1245; Q9UB1246; Q9UB1247; Q9UB1248; Q9UB1249; Q9UB1250;
AC Q9UB1251; Q9UB1252; Q9UB1253; Q9UB1254; Q9UB1255; Q9UB1256;
AC Q9UB1257; Q9UB1258; Q9UB1259; Q9UB1260; Q9UB1261; Q9UB1262;
AC Q9UB1263; Q9UB1264; Q9UB1265; Q9UB1266; Q9UB1267; Q9UB1268;
AC Q9UB1269; Q9UB1270; Q9UB1271; Q9UB1272; Q9UB1273; Q9UB1274;
AC Q9UB1275; Q9UB1276; Q9UB1277; Q9UB1278; Q9UB1279; Q9UB1280;
AC Q9UB1281; Q9UB1282; Q9UB1283; Q9UB1284; Q9UB1285; Q9UB1286;
AC Q9UB1287; Q9UB1288; Q9UB1289; Q9UB1290; Q9UB1291; Q9UB1292;
AC Q9UB1293; Q9UB1294; Q9UB1295; Q9UB1296; Q9UB1297; Q9UB1298;
AC Q9UB1299; Q9UB1300; Q9UB1301; Q9UB1302; Q9UB1303; Q9UB1304;
AC Q9UB1305; Q9UB1306; Q9UB1307; Q9UB1308; Q9UB1309; Q9UB1310;
AC Q9UB1311; Q9UB1312; Q9UB1313; Q9UB1314; Q9UB1315; Q9UB1316;
AC Q9UB1317; Q9UB1318; Q9UB1319; Q9UB1320; Q9UB1321; Q9UB1322;
AC Q9UB1323; Q9UB1324; Q9UB1325; Q9UB1326; Q9UB1327; Q9UB1328;
AC Q9UB1329; Q9UB1330; Q9UB1331; Q9UB1332; Q9UB1333; Q9UB1334;
AC Q9UB1335; Q9UB1336; Q9UB1337; Q9UB1338; Q9UB1339; Q9UB1340;
AC Q9UB1341; Q9UB1342; Q9UB1343; Q9UB1344; Q9UB1345; Q9UB1346;
AC Q9UB1347; Q9UB1348; Q9UB1349; Q9UB1350; Q9UB1351; Q9UB1352;
AC Q9UB1353; Q9UB1354; Q9UB1355; Q9UB1356; Q9UB1357; Q9UB1358;
AC Q9UB1359; Q9UB1360; Q9UB1361; Q9UB1362; Q9UB1363; Q9UB1364;
AC Q9UB1365; Q9UB1366; Q9UB1367; Q9UB1368; Q9UB1369; Q9UB1370;
AC Q9UB1371; Q9UB1372; Q9UB1373; Q9UB1374; Q9UB1375; Q9UB1376;
AC Q9UB1377; Q9UB1378; Q9UB1379; Q9UB1380; Q9UB1381; Q9UB1382;
AC Q9UB1383; Q9UB1384; Q9UB1385; Q9UB1386; Q9UB1387; Q9UB1388;
AC Q9UB1389; Q9UB1390; Q9UB1391; Q9UB1392; Q9UB1393; Q9UB1394;
AC Q9UB1395; Q9UB1396; Q9UB1397; Q9UB1398; Q9UB1399; Q9UB1400;
AC Q9UB1401; Q9UB1402; Q9UB1403; Q9UB1404; Q9UB1405; Q9UB1406;
AC Q9UB1407; Q9UB1408; Q9UB1409; Q9UB1410; Q9UB1411; Q9UB1412;
AC Q9UB1413; Q9UB1414; Q9UB1415; Q9UB1416; Q9UB1417; Q9UB1418;
AC Q9UB1419; Q9UB1420; Q9UB1421; Q9UB1422; Q9UB1423; Q9UB1424;
AC Q9UB1425; Q9UB1426; Q9UB1427; Q9UB1428; Q9UB1429; Q9UB1430;
AC Q9UB1431; Q9UB1432; Q9UB1433; Q9UB1434; Q9UB1435; Q9UB1436;
AC Q9UB1437; Q9UB1438; Q9UB1439; Q9UB1440; Q9UB1441; Q9UB1442;
AC Q9UB1443; Q9UB1444; Q9UB1445; Q9UB1446; Q9UB1447; Q9UB1448;
AC Q9UB1449; Q9UB1450; Q9UB1451; Q9UB1452; Q9UB1453; Q9UB1454;
AC Q9UB1455; Q9UB1456; Q9UB1457; Q9UB1458; Q9UB1459; Q9UB1460
```


DR 736 RNTMFGHTNSVNHCRFSPDDKLLASCADGTLKLDATSANF-----RKSINVKQFF 787
 QY 254 LNLNEDPNADLEFOODEGNIY-CYNWYGDG-RIMIGFSCGHFYVISTHTEGICEI--- 308
 DB 788 LNLNEDP-----QEDMEVYKCCSMASDAGARIWA-AKNHIFLEDHITSGLGEIHG 838
 QY 309 -----FOARNHKDNLSTAVSQTLNKVAATCGDNCIKIODLVLDKMDVYLNDEN 359
 DB 839 HHSITQYCDSPQNH---LAVVALSQ-----YCYE-----LMTDSRS 873
 QY 360 KGL---GTLNW-----TDDGOLLALSTORGSLHVFILFPLTLDGACSTRIVATLSLEY 410
 DB 874 KYADCRGHLNWGHVGFSPDGSSFLTSDQITRLMETK---KYCKNSAVMLKQEVNV 928
 QY 411 T-----VANVEGELPTVSDVEPEPFAVGLVHLAVGNRRAMFVY 452
 DB 929 VQGENEVVLANVDHRLRLDINGRTGQIDYLEAQVSCCCLSPHIOYIAFGENGCA---- 984
 QY 453 LGENAVKLLKDMYVIGYASICLHSDYAAALFEGKVQULHLESLIDAOERETRLFPAY 512
 DB 985 -----TEILELVNRRIFQSRFOHKTVMHIOFTADEKTLISSSDAEIQVMNQ 1033
 QY 513 DDKC-----RLCHALTSDELIVGTGTVGVQYFIEDMOFVNDYRHPVSVK 559
 DB 1034 LDKCIFLNGHOETVYKDFRLKNSRLLSWSEDT-----VKYWNITGK-----EK 1079
 QY 560 IFPPDNGTRIVFIDEKSDGFVYCPVNDATYEIPEFSPTIKGVLENMPMD 609
 DB 1080 DFVCHQGVTL-----SC---DISHDTRKFSSTADTAIMSRD 1115

RESULT 13

DR 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DE 16-OCT-2001 (Rel. 40, Last annotation update)
 WD-repeat protein 1 (Actin interacting protein 1) (XAIP1).
 GN WDRL OR AIP1.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
 OC Xenopodidae; Xenopus.
 NC NCB1_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99230343; PubMed=10212149;
 RA Okada K., Obinata T., Abe H.;
 RT "XAIP1: a xenopus homologue of yeast actin interacting protein 1
 with ADF/cofilin family proteins.";
 RT J. Cell Sci. 112:1553-1565(1999).
 CC -1- FUNCTION: INDUCES DISASSEMBLY OF ACTIN FILAMENTS IN CONJUNCTION
 WITH ADF/COPILIN FAMILY PROTEINS.
 CC -1- SIMILARITY: CONTAINS 11 WD REPEATS (TRP-ASP DOMAINS).
 CC -1- SIMILARITY: BELONGS TO THE AIP1 FAMILY OF WD-REPEAT PROTEINS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).
 CC
 DR EMBL, AF124140; AAD22062.1; -
 DR InterPro; IPR001680; WD40.
 DR Pfam; PF004400; WD40. 9.
 DR PRINTS; PRO0320; GPROTEINRPT.
 DR SMART; SM00320; WD40. 11.
 DR PROSITE; PS00676; WD_REPEATS_1; 1.
 DR PROSITE; PS50082; WD_REPEATS_2; 4.

DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
 KW Actin-binding; Cytoskeleton; Repeat; WD repeat.
 FT REPEAT 56 95 WD 1.
 FT REPEAT 100 143 WD 2.
 FT REPEAT 145 184 WD 3.
 FT REPEAT 187 226 WD 4.
 FT REPEAT 232 271 WD 5.
 FT REPEAT 318 359 WD 6.
 FT REPEAT 363 401 WD 7.
 FT REPEAT 436 482 WD 8.
 FT REPEAT 487 526 WD 9.
 FT REPEAT 530 569 WD 10.
 FT REPEAT 575 607 WD 11.
 SQ SEQUENCE 608 AA; 66092 MW; C8C0F8A47A4759F0 CRC64;

Query Match 2.0%; Score 147; DB 1; Length 608;
 Best Local Similarity 19.1%; Pred. No. 0.1; 212; Indels 182; Gaps 24;
 Matches 112; Conservative 81; Mismatches 212; Indels 182; Gaps 24;

QY 40 EKRRIFSLLEKTWLGAPIQFAMQKTSQNYLAVTGADYIK-----IFDRH----- 84
 DB 4 ELKKYFASLPQMERGVAKTLADGPKGNFLYINGSVIIRINDPALADIYTEHHAQAV 63
 QY 85 -----GKRSEINL-----PGNCVAMDWDKGDVLAIVA-- 113
 DB 64 ARYASGFIYASGDSGKRLIMDTQKEHLKYEYQFAGKIKDIATWEDSKRIAVVGE 123
 QY 114 -EKSSCIYIMDANTN-----KTSQDLN----- 134
 DB 124 RKRFASVFLMDTGSSVGGEITGNIKVINSVDIKOTRPRYLVTGSDNCAFLGPPPKRF 183
 QY 135 GMRDMSF---LLMSKVSFLAVGTGKGLXLYNHQTSKRIPLVLS---KHTKRITCGGN 188
 DB 184 TMSDSRFVNCVRFSPDGSKLASAGADGQIFLYDKRTGKGVSLGSAHOGGIYAVGMS 243
 QY 189 AENLXALGE-DKMITVSQSGDTIRQT-QVRSEPKNQFFLMKDDRTSAAESWISVYL 246
 DB 244 PDGTOILLASAGSKTKYKIMDVANSVTFTEFNLGSDVLDQGLGIMOKD-----YLLSVSL 297
 QY 247 GKTLFFPLNLNPPADLEFOODEGNIYCY---HWYGGRIIMFSCGHFVISTHGE 303
 DB 298 -SGYINIDKKNPAPRLV-IKGNKSTQCMVTWNNSDGRSTIYTSHGHIWMDAETGE 355
 QY 304 LGQELFOARNHKDNLSTAVSQTLNKVAATCG-DNCIKIODLV----- 344
 DB 356 --NDFTFGKHTNQVSSMDL-DGCGQLITCSMDLTRLSLSKSYSSSESVMQVQKC 412
 QY 345 -----DLKDMYVILNDEE-----NKGIGTSL-WTDDGOLLAL 376
 DB 413 VAVSGGVVTVICQIQLVLLKDKKKVFADISLDYEPKVAALHKGSTVAVAGGADKRVHL 472
 QY 377 SGORSLHVFILKPLILDGACSTRIVATLSLEYVAVNVEGELPTVSDVEPEPFAVG 436
 DB 473 STQGSLSLDECKTLIPAKAVYDLDVLSHGCAFLAVDAKV-----VT 514
 QY 437 LYHLAVGNRRAMFVYLGENA-----VKRLKDMYVIGT-----VASICL 475
 DB 515 VFSVADGVSEKNSY--GHAAKALSVAMSPDNEHFASSGMGMGVCCL 559
 RESULT 14
 DR WDRL_CHICK
 AC 093277;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DE 16-OCT-2001 (Rel. 40, Last annotation update)
 WD-repeat protein 1 (Actin interacting protein 1).
 GN WDRL.
 OS Gallus gallus (chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;

```

OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99156857; PubMed=10036186;
RA Adler H.J., Wlinski R.S., Gong T.-W.L., Lomax M.I.;
RT "A gene upregulated in the acoustically damaged chick basilar papilla
  encodes a novel WD40 repeat protein."
RL Genomics 56:59-69(1999).
CC -1- FUNCTION: INDICES DISASSEMBLY OF ACTIN FILAMENTS IN CONJUNCTION
  WITH ADF/COPFILIN FAMILY PROTEINS (BY SIMILARITY).
CC -1- SIMILARITY: CONTAINS 11 WD REPEATS (TRP-ASP DOMAINS).
CC -1- SIMILARITY: BELONGS TO THE APLI FAMILY OF WD-REPEAT PROTEINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF020054; AAD05042.1; -
DR InterPro: IPR001680; WD40.
DR Pfam: PF00400; WD40; 10.
DR PRINTS: PR00320; GPROTEINBRPT.
DR SMART: SM00320; WD40; 11.
DR PROSITE: PS00678; WD_REPEATS_1; 1.
DR PROSITE: PS50082; WD_REPEATS_2; 5.
DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
DR Actin-binding; Cytoskeleton; Repeat; WD repeat.
KW REPEAT
FT REPEAT 58 97 WD 1.
FT REPEAT 102 145 WD 2.
FT REPEAT 146 186 WD 3.
FT REPEAT 189 228 WD 4.
FT REPEAT 234 273 WD 5.
FT REPEAT 320 361 WD 6.
FT REPEAT 365 403 WD 7.
FT REPEAT 445 484 WD 8.
FT REPEAT 489 528 WD 9.
FT REPEAT 532 571 WD 10.
FT REPEAT 576 608 WD 11.
SQ SEQUENCE 609 AA; 66566 MW; 1DE89BE431DEE32A CRC64;

Query Match 2.0%; Score 143.5; DB 1; Length 609;
Best Local Similarity 19.9%; Pred. No. 0.17;
Matches 125; Conservative 92; Mismatches 238; Indels 173; Gaps 32;

OY 65 SGNVLAVTGADYIVKIFD---RHGQRSEINLPNCVAMDMDDKGDVLAIVIA--EKSS 117
DB 71 SGNVLAISGVSGSKRLIMDTTQKREHLKYEQFPAGKIKDLAMVEDSKRIAVVEGRKRF 130
OY 118 CIYLDAMNT-----NK-----TSQLDN-----GMDQMSFLL-- 144
DB 131 AVFLAMSGSVGEITGNKINSVDIKQTRYLATGSDNCAAFEGPPEKFTLSDH 190
OY 145 -----ASKVGSFLAVGVKGNLKIYNQTSRKIPVIG--KHKRKITCGGMNENLX 193
DB 191 TRFVNCVRSSPDGNRPRAVSADQIFLYDGKTEKVCALGGKAHDGCIYAISMSPSSQ 250
OY 194 ALGGE-DKMITVSNQEGDTIRQT-QVRSSEPKNMQFLMKMDRTSAESMISVLGKKTLL 251
DB 251 LLSASDCKTAKIKWDGANSVSTFNMGSNVLDDQLGCLMQKDH-----LLSLSL-SCYI 303
OY 252 FFLNLNEPDNPADLEPQDFGNIVCYINW---GDGRIMIGFSCGHEVYVSTHTGELQET 308
DB 304 NYFDKNNPKPLRY-IKGHSKSIQCLTVHKNKGSKSYIYSGSNDGHNWYWDSDGE--NDG 360
OY 309 FOARNHKNDNTSTAVSGTQTKKVAATCGDNCKIKDLDVLAQMYVILNDEENKKGITLSWT 368
DB 361 FSGKGHTNOVSRAVDE-MQDLVTC-----SMDD-----TVRYT 393

```

```

OY 369 DDCQLLALSTORGSLLHFLKRLPILG-DACSTRIAYLTSLLEVTVANPVGEGLPIYVSVD 427
DB 394 -----LSKRDYSGDA-----YKMD 409
OY 428 VEPNFAVAG---LYHLANGMNNRAMEFYVLEGNAAVK--KIKMEYIGTVASICLH---SD 478
DB 410 VQFKCLAVGGGTYVVICIQ-----IYLMKKKKCFADIDGLY--EPEAVAVHPOGGS 461
OY 479 YAAALFEGYQOLHISEELIDAOEERTRLFPVADKCRILCHALYSDFLYGTDRGVQ 538
DB 462 VAVGGTGDGNVRLSIQTSLSKSDKTLKAKGPVTD-----LAISHDGAFLAVCDANKVVT 516
OY 539 YFYEDMQFVND--YRHPVSVKRIFFDPNCTRLVFLDEKSDGEVYC-PVND--ATYEIPD 593
DB 517 VFSVPGGVYEHNFYGHNAVVCIAISPDEH--FASGMDMVMYVTVSDPETRIKIPD 574
OY 594 FSPF--IKGVLEWNPMDKGVFIAYDDDD 619
DB 575 AHRLLHVSGLAW-----LDEHTLVTTSHD 598

RESULT 15
CLH_CAEEL STANDARD; PRT; 1681 AA.
AC P34574;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Probable clathrin heavy chain.
GN T2065.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Berkis M., Smith A.;
RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: CLATHRIN IS THE MAJOR PROTEIN OF THE POLYHEDRAL COAT OF
  COATED PITS AND VESICLES (BY SIMILARITY).
CC -1- SUBUNIT: CLATHRIN TRISKELIONS, COMPOSED OF 3 HEAVY CHAINS AND 3
  LIGHT CHAINS, ARE THE BASIC SUBUNITS OF THE CLATHRIN COAT
  (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC FACE OF COATED PITS AND
  VESICLES (BY SIMILARITY).
CC THE TRISKELION, THIS REGION CONTAINS THE TRIMERIZATION DOMAIN AND
  THE LIGHT-CHAIN BINDING DOMAIN INVOLVED IN THE ASSEMBLY OF THE
  CLATHRIN LATTICE.
CC -1- SIMILARITY: BELONGS TO THE CLATHRIN HEAVY CHAIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z30423; CAA83003.1; -
DR PIR: S42369; S42369.
DR HSSP: P14442; 1BPO.
DR WormRep: T2065.1; CE00480.
DR InterPro: IPR001473; Clathrin_propel.
DR InterPro: IPR000547; Clathrin_repeat.
DR Pfam: PF01394; Clathrin_propel; 7.
DR Pfam: PF00637; Clathrin_repeat; 7.
DR SMART: SM00299; CLH; 7.
KW Hypothetical protein: Coated pits.
SQ SEQUENCE 1681 AA; 191541 MW; 44D15C61339009D9 CRC64;

```


THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 21, 2002, 14:49:40 ; Search time 50.02 Seconds
(without alignments)
4779.668 Million cell updates/sec

Title: US-09-729-653-2
Perfect score: 7257
Sequence: 1 HSLIGRCSRLGIDGNNAVC.....AAQLKTSIDCTQYLRTHEEL 1382

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues
Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%
Listing first 45 summaries

Database :

SPREMBL_19:*
1: sp.archaea:*
2: sp.bacteria:*
3: sp.fungi:*
4: sp.human:*
5: sp.invertebrate:*
6: sp.mammal:*
7: sp.mhc:*
8: sp.organelle:*
9: sp.phage:*
10: sp.plant:*
11: sp rodent:*
12: sp.virus:*
13: sp.vertbrate:*
14: sp.unclassified:*
15: sp.rvivirus:*
16: sp.bacteriap:*
17: sp.archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4700	64.8	905	4	Q9HCD4
2	2780	38.3	1366	5	Q9V8W2
3	2354	32.4	481	4	Q9H5S0
4	1157	15.9	519	5	Q46018
5	924	12.7	634	5	Q46020
6	289	4.0	1205	5	Q9W097
7	289	4.0	1749	11	Q9UKU3
8	271.5	3.7	1242	4	Q9HBG6
9	270.5	3.7	1758	5	Q22830
10	268.5	3.7	1292	4	Q9HBG5
11	263	3.6	1198	4	Q9UF80
12	263	3.6	1241	4	Q9HAT9
13	256.5	3.5	1462	4	Q60332
14	251.5	3.5	1653	4	Q9UG01
15	248.5	3.4	1462	4	Q96RY7
16	246	3.4	1131	4	Q9BTR9

17	246	3.4	1160	4	Q9BTR4	Q9BTR4 homo sapien
18	235.5	3.2	993	4	Q9NV68	Q9NV68 homo sapien
19	235	3.2	1443	5	Q9VPR0	Q9VPR0 drosophila
20	223.5	3.1	955	4	Q9H7P3	Q9H7P3 homo sapien
21	218	3.0	764	4	Q9P2H3	Q9P2H3 homo sapien
22	216.5	3.0	864	4	Q9H9Z0	Q9H9Z0 homo sapien
23	215.5	3.0	705	4	Q96KK0	Q96KK0 homo sapien
24	209.5	2.9	1307	5	P90757	P90757 caenorhabd
25	206	2.8	821	4	Q9NRM6	Q9NRM6 homo sapien
26	190.5	2.6	1119	5	Q18859	Q18859 caenorhabd
27	190	2.6	766	4	Q9P2L0	Q9P2L0 homo sapien
28	185	2.5	1090	4	Q9ULP1	Q9ULP1 homo sapien
29	183.5	2.5	588	4	Q96KK1	Q96KK1 homo sapien
30	183.5	2.5	656	4	Q9UG52	Q9UG52 homo sapien
31	180	2.5	601	11	Q9UYV4	Q9UYV4 mus musculu
32	178	2.5	1047	5	Q45401	Q45401 caenorhabd
33	174.5	2.4	968	5	Q9V5J6	Q9V5J6 drosophila
34	171.5	2.4	1319	10	Q9FNA4	Q9FNA4 arabidopsis
35	170	2.3	887	5	Q9W040	Q9W040 drosophila
36	164	2.3	472	10	Q9FEV8	Q9FEV8 arabidopsis
37	164	2.3	493	4	Q9U196	Q9U196 homo sapien
38	164	2.3	493	4	Q75869	Q75869 homo sapien
39	164	2.3	504	4	Q9ULH8	Q9ULH8 homo sapien
40	163	2.2	496	4	Q9UM11	Q9UM11 homo sapien
41	161.5	2.2	558	4	Q9UFJ9	Q9UFJ9 homo sapien
42	161.5	2.2	911	10	Q9LW87	Q9LW87 arabidopsis
43	161.5	2.2	1223	4	Q9URF5	Q9URF5 homo sapien
44	161	2.2	493	11	Q9RIK5	Q9RIK5 mus musculu
45	161	2.2	493	13	Q42585	Q42585 xenopus lae

ALIGNMENTS

RESULT 1
ID Q9HCD4 PRELIMINARY; PRT; 905 AA.
AC Q9HCD4:
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE K1A1638 PROTEIN (FRAGMENT).
GN K1A1638.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20450683; PubMed=10997877;
RA Nagase T., Kikuno R., Nakayama M., Hirosewa M., Ohara O.;
RT Prediction of the coding sequences of unidentified human genes.
RT XVIII. The complete sequences of 100 new cDNA clones from brain which
RT code for large proteins in vitro."
RL DNA Res. 7:273-281(2000).
DR EMBL: AB046858; BAB13464.1;
DR InterPro: IPR000547; Clathrin_repeat.
DR SMART: SM00299; CLH; 1.
FT NON_TER
SO SEQUENCE 905 AA; 102450 MW; FD370996F00FE1BF CRC64;

Query Match 64.88; Score 4700; DB 4; Length 905;
Best Local Similarity 99.88; Pred. No. 0;
Matches 903; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 375 ALSTQPSGLVFLTKPLTGACSTRVATLSLETVANPVEGELPTVSVDPEPFVA 434
DB 1 ALSTQPSGLVFLTKPLTGACSTRVATLSLETVANPVEGELPTVSVDPEPFVA 60
QY 435 VGLYHLAVGNMNRAMPVYLGENDAVKRLKDEYLGTVASICLHSDYAAALFEKGKYLHIE 494
DB 61 VGLYHLAVGNMNRAMPVYLGENDAVKRLKDEYLGTVASICLHSDYAAALFEKGKYLHIE 120

```

OY 495 SEILDAOEERERLEPAVDKCRILCHALTSPFLYIGTGTGVVQYFIEDMOQFVNDYRHP 554
    |||||
Db 121 SEILDAOEERERLEPAVDKCRILCHALTSPFLYIGTGTGVVQYFIEDMOQFVNDYRHP 180
OY 555 VSVKRIPEPNCNTRIVFIDKSDGFYCCVNDATYEIPFSPITKGVLMENPMKGVFI 614
    |||||
Db 181 VSVKRIPEPNCNTRIVFIDKSDGFYCCVNDATYEIPFSPITKGVLMENPMKGVFI 240
OY 615 AYDDKRVYTVFHKDTIOGAKVILAGSTKVPFAHKPRLLYNSELNCOGOSKVNNTYST 674
    |||||
Db 241 AYDDKRVYTVFHKDTIOGAKVILAGSTKVPFAHKPRLLYNSELNCOGOSKVNNTYST 300
OY 675 HGFLESLKDXGDELPRMPLAHLMLKRESDAWEMCRILNDEAAMNELARACHHEVEVERA 734
    |||||
Db 301 HGFLESLKDXGDELPRMPLAHLMLKRESDAWEMCRILNDEAAMNELARACHHEVEVERA 360
OY 735 IRVYRIRIGVNGIVMSLEOIKGIEDYINLAGHLMFNDYNLADODYLASSCPILALEMR 794
    |||||
Db 361 IRVYRIRIGVNGIVMSLEOIKGIEDYINLAGHLMFNDYNLADODYLASSCPILALEMR 420
OY 795 DLQWDSALQIAKHLAPDQIPFISKEVATOLEFAGDYVNALAHYEKGITGDKNEHDEACT 854
    |||||
Db 421 DLQWDSALQIAKHLAPDQIPFISKEVATOLEFAGDYVNALAHYEKGITGDKNEHDEACT 480
OY 855 AGVAMSTIRMGDIRRGVNGVNAKHPBRVILKRCOGATLENNKOFSEAAOLYERKLYYDKAAS 914
    |||||
Db 481 AGVAMSTIRMGDIRRGVNGVNAKHPBRVILKRCOGATLENNKOFSEAAOLYERKLYYDKAAS 540
OY 915 VIIRSKNNAKVGDLLPHVSSPKIHLQYAKAKADGRYKAAVAVAYENAKOMOSVIRIYDTH 974
    |||||
Db 541 VIIRSKNNAKVGDLLPHVSSPKIHLQYAKAKADGRYKAAVAVAYENAKOMOSVIRIYDTH 600
OY 975 LNPEKAVNIYRETOSLDGAKKAVANFLOLDGYSAIOFLVNSKCNNEAFTLAQOHNKME 1034
    |||||
Db 601 LNPEKAVNIYRETOSLDGAKKAVANFLOLDGYSAIOFLVNSKCNNEAFTLAQOHNKME 660
OY 1035 IYADITIGSEDTNEDYOSATLFEBEKRTLOAGKFFLLCGOVSRAKHLKPPSESDNVA 1094
    |||||
Db 661 IYADITIGSEDTNEDYOSATLFEBEKRTLOAGKFFLLCGOVSRAKHLKPPSESDNVA 720
OY 1095 IEMALETVGOAKDELITNOLIDHLLGENDGMPKDAKYLFRILYMAKQYREAAQTATIIAR 1154
    |||||
Db 721 IEMALETVGOAKDELITNOLIDHLLGENDGMPKDAKYLFRILYMAKQYREAAQTATIIAR 780
OY 1155 EEOGAGNTRNADVLFSMTAEKLSOKIKIPSEMATNMLTSHSYILVKIHVKNGDHMKGAR 1214
    |||||
Db 781 EEOGAGNTRNADVLFSMTAEKLSOKIKIPSEMATNMLTSHSYILVKIHVKNGDHMKGAR 840
OY 1215 MLIRVANNISKEPFSHIVPILSTVTECHRAGLKNASFAALMLRPEYRSKIDAKYKKKI 1274
    |||||
Db 841 MLIRVANNISKEPFSHIVPILSTVTECHRAGLKNASFAALMLRPEYRSKIDAKYKKKI 900
OY 1275 EGMVR 1279
    ||||
Db 901 EGMVR 905

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Burton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Branton R.C., Rogers Y.-H.C., Blazer R.G., Champagne M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Abmayan A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Baller R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Broststein P., Brothier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cieslewski S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo S., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferlita S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostlin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
RA Jaitani M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazozo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheibel F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Svirker E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Sviderskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
CC -I- SIMILARITY: CONTAINS 6 WD REPEATS (TRP-ASP DOMAINS).
DR EMBL: AE003795; AAF57545.1; -
DR FlyBase: FBgn0034452; CG11237.
DR InterPro: IPR001680; WD40.
DR Pfam: PF00400; WD40; 6.
DR SMART: SM00320; WD40; 4.
DR Repeat: MD repeat.
SQ SEQUENCE 1366 AA; 153505 MW; E9EBDC9C61AFBBBD CRC64;

```

Query Match 38.3%, Score 2780; DB 5; Length 1366;
 Best Local Similarity 40.8%; Pred. No. 1.2e-182;
 Matches 553; Conservative 286; Mismatches 468; Indels 50; Gaps 14;

```

OY 57 IOFAMOKTSGNLAATGADYIYKIFDRHGOKRSEINLPGNCVAMDMDGDVLAIAKRS 116
    ||||
Db 19 YEFIWOKA---LATTGTGDSVALYNROGLVORITLISGLSGFAMDDEGDLGITSGS 75
OY 117 SCIYLMADANTNKTSQLDGMGRMDOMSFMLMSKVGSLAVGTAGVGNLXINHGOT-SKRIPL 175
    ||||
Db 76 PRITLMDYNSQEKISYEVGLRDLPLCLILMSKQOQLAVGTGNGALNATYINHSKGKPTVYL 135
OY 176 GKHTKRITCGCWNALXALGEGDKMTVSNQEGDTIROTVRESEPPXNMQFFLMKMDRT 235
    ||||
Db 136 GKHSKRITCGAMSAQNLLATLGESEKFSLSNEDGDTVAVYVLRDAPDTMYFAEMANDERI 195
OY 236 SAAESMISVNLGKRTFLPLNLNPNPNADLEFQDFGNIVCNWYGDGRIMIGFSCGFY 295
    ||||
Db 196 -AGDAISMIIIGKRTFLFYLPENPELGEFSQSRGSLMDHKMGWGIILLGFGNHV 254
OY 296 VISTHTGEOEIFOARHKKNLNTSIAVSQTLNKAATGCDKICRODVLDKMVIINL 355
    ||||

```

```

Db 255 AISTHPRVGOELMOVKNHNSDLTGLAYCPTLDIVASCDDSGIKHSITNLOJETERITV 314
QY 356 DEENKGLTSMWTDGOLLALSTORGSLHVELKRLPLDAGSTRATAYLSLELVY--- 412
Db 315 -PBAAGVOMIDWSPDQOLLAVTNNHGVYIYVTKPLHLEAVSPRLVLSLAEVSIYV 373
QY 413 ANVEGELPTTVADVDEPNVAAGLYHLAVGNMRAFFVYLGEN---AVKELDMETIGT 469
Db 374 APKRTSLPRRLPLEBEPPTMAVGFNFATGIEKHWFYDGLSGSEEPRLSERPPRS 433
QY 470 VASICLHSDYAAALFEKVOHLIESEILDQOERET---RLFPVADCKRILCHALTS 526
Db 434 VESKINADYCALCPQLLQAIADNPCKDLQAVFETPALNPMSDVAIVITCFALSOE 493
QY 527 FLITGDTGVVOYFYIEDNGFVNDYRHPVSKFIIFDPNCTRLVFIDEKSDGVYVCND 586
Db 494 LLLFADIDGLHVFYSLEKMSCTIYRHSWIGIROLFMDIESTKFIIFDHSQGVFLPVE 553
QY 587 ATYEIPDFPTTIGVLEMMMPDKGVFIADVDDKVYTYVFNKDTIOGAKVILGSTRKVP 646
Db 554 EALLIDIPKOCGLAMD--LTQPNIFISDARIIVTHVVRHSVQGTHTLVMGESKLP 611
QY 647 AHKPLLLYNGELTQTOGKVNNTYSTHGLSNLKOXGPDELRLPHANLMLKRSDDAW 706
Db 612 GQPLLLCGEMALHIDGQYATQSLSTH--VYVNSQANLQML---LKLRYNDEAY 665
QY 707 EMKRIINDEAANMLRACLHMEFATVYRIGVNGVMSLEQIGTIEDYNLLAGHL 766
Db 666 KLCCKOMHOSAMREFEQALSDLEPDLATRAYQLGSAALVNLSEIRYEDLMLGCC 725
QY 767 AMFTNDYNLAODLYLASSCPALLENMRDLQHNDSALQALAKHLAPDQIPRESKAYOLE 826
Db 726 CTLAQYDQAKENHMLGVYRAALDLCDRLQMDQALLLHKNKDPQVRLTAREVYQOLE 785
QY 827 FADYVNAALAHYKGTGD-----NKHDEACLAGVQOMSTRMDIR 868
Db 786 FGNVYDADALHYEKYKEDINSKETETDALMSPEYEHVRLCKGSIARTSIRADFR 845
QY 869 RGNVALKHSRYLKRDCGAILENMKOFSEAOQLYEKGILYDKAASYIASKMAKYGDL 928
Db 846 RGVQYVLEEDQOLDFDCAELTAVGHLETAAGLYERGFYDECGHYTLAKMKNKANNI 905
QY 929 LPHVSSPKIHLQYAKAKADAGRYKAVAVAYENAKOMOSVIRIYLDHNLNPEKAVNIRET 988
Db 906 LPRVSKTKLHAAYAKAKENGHVEAIRSYRIGADDACRYITLDHLCODHNAASEIYLES 965
QY 989 OSUDGAKMVARFLOLDIGSAYIQFLVMSKCNNEAFTLAOQHNMKEIYADIGS-EDTNN 1047
Db 966 RSMDSAKLLAKFYOKIGDVQALQFLVTCGVEEAFALQHRNKLRLRHGELLEERYENAKS 1025
QY 1048 EDVOSIALYEEGRKRYLOACKFLLCGOVSRAKHLKCP--SSEDNVATLMAIETVGOA 1105
Db 1026 SDVLALAHFEGERKRYTLACKYFLAREFYKALRFLKASAFNNEBOVSJSLAIDCVATS 1085
QY 1106 KDELLTNQILDHLLGENDGAPKDAKYLFLRYMALKOYREAAQTALITAREOSAGNYRNA 1165
Db 1086 NNEQALATQLEIFLGEVDGVPKPRYFLRYMARKNHYKDAKATAVILANEOQLAGNKSA 1145
QY 1166 HDVLFKSYAELKSOIKIRBEMATNMLHSYLLVTKHYVNGDHMKCARMLIRYANNISK 1225
Db 1146 RDLVSYVOELRRNNLSVTREMRHOFILRLHYLVRILHGNHLLAKLALVOYACISO 1305
QY 1226 FPHSHVPLSTVYIECHRAALKNSAFSFAMLMRPREYRSIDAKYKKIKGVMRRP-- 1282
Db 1206 FPHHITPILSTVYIECHRAALKNSAFSFAMLMRPREYRSIDAKYKKIKGVMRRP-- 1265
QY 1283 ---ISEIEATTPCPFCFLPLPECELCPCKKNSIPYCIATGRHMLKDMVTPVCHDFPA 1339
Db 1266 KOLRDEIDGEMCPICDSNLANNEVTCYSCKTLPLCIATGQILIKQMLTSCQOCPFLC 1325
QY 1340 LYSELKATMLNTSTCPKCSERLNAOLKISDCTQYL 1376
Db 1326 FRAEMENILSENGECPMCGENVAPEOLLDAVEDIRPYI 1362

```

```

RESULT 3
AC 09H550 PRELIMINARY; PRT; 481 AA.
ID 09H550;
AC 09H550;
DT 01-MAR-2001 (TREMELREL, 16, Created)
DT 01-MAR-2001 (TREMELREL, 16, Last sequence update)
DT 01-DEC-2001 (TREMELREL, 19, Last annotation update)
DE CDNA: FLJ23127 FIS, CLONE LMG08350.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LUNG;
RA Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T., Matsumura K.,
RA Nakajima Y., Mizuno T., Morinaga M., Tanigami A., Fujiwara T., Oro T.,
RA Yamada K., Fujii Y., Ozaki K., Hiroo M., Ohmori Y., Ota T., Suzuki Y.,
RA Odayashi M., Nishi T., Shibahara T., Tanaka T., Nakamura Y.,
RA Isogai T., Sugano S.;
RT NEDO human cDNA sequencing project.;
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 3 WD REPEATS (TRP-ASP DOMAINS).
DR EMBL: AK026780; BAB1550.1;
DR InterPro: IPR001680; WD40.
DR Pfam: PF00440; WD40; 3.
DR SMART: SM00320; WD40; 3.
KW Repeat; WD repeat.
SQ SEQUENCE 481 AA; 53542 MW; B1C7AE096581F02 CRC64;

```

```

Query Match 32.4%; Score 2354; DB 4; Length 481;
Best Local Similarity 99.3%; Pred. No. 5,5e-154;
Matches 449; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 41 MKRIFSLLEKTMWLGAPIOFAMOKTSGNYLAVTGADYVYKIFDRHGRSEINLPGNCVAM 100
Db 1 MKRIFSLLEKTMWLGAPIOFAMOKTSGNYLAVTGADYVYKIFDRHGRSEINLPGNCVAM 60

QY 101 DMDKGDVLAIVAEKSSCIYLDANTNKTSQLDNGMRDMSFLMSKVSFLAVGTVKN 160
Db 61 DMDKGDVLAIVAEKSSCIYLDANTNKTSQLDNGMRDMSFLMSKVSFLAVGTVKN 120

QY 161 LAYINROTSKRTIYVGNHTRITTCGCMNAENLALGDEKMTIYVSNQEGTITQTVRSE 220
Db 121 LAYINROTSKRTIYVGNHTRITTCGCMNAENLALGDEKMTIYVSNQEGTITQTVRSE 180

QY 221 PXMNPOFLMKMDRTSAEEMISVYLGKKTFLFNLNEPNDPDLFQOQFGNIVCYNWY 280
Db 181 PXMNPOFLMKMDRTSAEEMISVYLGKKTFLFNLNEPNDPDLFQOQFGNIVCYNWY 240

QY 281 GDRIMIGFSGHGFVYISTHTGELGQEIFQARNHKNLTSIAVSQTLNKVATCGDNCIKI 340
Db 241 GDRIMIGFSGHGFVYISTHTGELGQEIFQARNHKNLTSIAVSQTLNKVATCGDNCIKI 300

QY 341 ODVLDLKMVYIILNDENGGLSTWTDGQLLALSTORGSLHVELTKPLPLGDACSTR 400
Db 301 ODVLDLKMVYIILNDENGGLSTWTDGQLLALSTORGSLHVELTKPLPLGDACSTR 360

QY 401 IAYTSLLEVTVANPVGEGPLITYSVDVDEPNFAVGLYHLAVGNMRAFFVYLGENAVKK 460
Db 361 IAYTSLLEVTVANPVGEGPLITYSVDVDEPNFAVGLYHLAVGNMRAFFVYLGENAVKK 420

QY 461 LKDMETVGTVAISICLHSDYAAALFEKVOHL 492
Db 421 LKDMETVGTVAISICLHSDYAAALFEKVOHL 452

RESULT 4
ID 046018 PRELIMINARY; PRT; 519 AA.

```

AC 046018: 1
DT 01-JUN-1998 (TREMBlrel. 06, Created)
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE ZK520.1 PROTEIN.
GN ZK520.1
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_Taxid=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Steward C.A.;
RL Submitted (MAR-1997) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RT none;
RT "Genome sequence of the nematode C.elegans: A platform for
investigating biology."
RL Science 282:2012-2018(1998).
DR EMBL; Z92822; CAB07299.1;
DR PROSITE; PS0294; WD_REPEATS_REGION; 1.
SQ SEQUENCE 519 AA; 59152 MW; 38F80CA32D06E02D CRC64;

Query Match 15.9%; Score 1157; DB 5; Length 519;
Best Local Similarity 45.0%; Pred. No. 2.3e-71;
Matches 230; Conservative 114; Mismatches 157; Indels 10; Gaps 4;

OY 724 AGLHMEVFATRYVRIGVGVMSLEQIKGIEDYNLLAGHMTNTNYNAODIYLAS 783
DB 2 AALLSDVMAKIFREIDDAVALELEITEERNLLHAQYTLISRLDAEOLYLES 61
OY 784 SCPIALENRRDLQHSALQALAKHLAPDQIPFISKEVAIQLEFADYVNAALAHYEKG- 842
DB 62 SRPMELNNRRDLLEPRKALVLAETNPKREIPYLSKEVQOELELGDHNSLANYEKGVM 121
OY 843 -----TGKNKEHDEACLAGVAQKSTMGDIRGVNOLKHPRSVYLKRCGAILNNKOPS 897
DB 122 ENPQNLPELQEHNEICQSGIARMAIKTGDRLRGVOLAKOLEGRVVRKDCAILLEOMKQYT 181
OY 898 EAAQLEYKGLYDKAASVYIRSKNNKAVGDLPHVSSPIHLQYAKKADGKRAVVA 957
DB 182 EAAQLEYKGLYDKAASVYIRSKNNKAVGDLPHVSSPIHLQYAKKADGKRAVVA 957
OY 958 YENAKQWQSVIRIYDLHNNPEKAVNIVETQSLDGAKVAREFLOLGDYSAIOFLVMS 1017
DB 242 YETGRDYDQVRLPLDPLNDPEAVVRESRSIEGAKLVAKFVYKLDGYNISAIQFLVMS 301
OY 1018 KNNNEAFTLAQOHNNKEITADIIGSEDITNEDYQSIALFESEKRLQAGKFFLQGYTS 1077
DB 302 QCVQAEFFELAEKNNMAVREYAKAIEQHGNISQALE-LAEYNNRVNMFMAKFFYTOAGQYN 360
OY 1078 RLKHLKPCSESDNVAIEMATEVQAGKDELITNOLIDLHLENDGPKMDKYLFLRM 1137
DB 361 NAINLFLK--KDDDENCVALAVDCGKSKDKTLNKLVLFLGE-DGNVYKDAQFLRLV 417
OY 1138 ALKQYREAAQTAIIAREBQSGANTYRNADVLFSMYAEKLSOKIKIPSEMATNLMLHSY 1197
DB 418 GLGRKDAQGTAVVAQIHQAGKNYRIARDLFLQMHQGLREKMMRPLDMNNSLMAIHSY 477
OY 1198 ILVKIHKNGDHMKGRMLIRVANNISKFPSS 1228
DB 478 ILVKALINRKETLLAARLLIFTCGEIQRPT 508

RESULT 5
O46020 PRELIMINARY; PRT; 634 AA.
AC 046020;
DT 01-JUN-1998 (TREMBlrel. 06, Created)
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)

DE ZK520.3 PROTEIN.
GN ZK520.3
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_Taxid=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Steward C.A.;
RL Submitted (MAR-1997) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RT none;
RT "Genome sequence of the nematode C.elegans: A platform for
investigating biology."
RL Science 282:2012-2018(1998).
CC -1- SIMILARITY: CONTAINS 4 WD_REPEATS (TRP-ASP DOMAINS).
DR EMBL; Z92822; CAB07301.1;
DR InterPro; IPR001680; WD40.
DR Pfam; PF00400; WD40; 4.
DR SMART; SM00320; WD40; 3.
DR PROSITE; PS0294; WD_REPEATS_REGION; 1.
KW Repeat; WD repeat.
SQ SEQUENCE 634 AA; 69905 MW; F35637A84B16930A CRC64;

Query Match 12.7%; Score 924; DB 5; Length 634;
Best Local Similarity 30.4%; Pred. No. 3.7e-55;
Matches 204; Conservative 130; Mismatches 239; Indels 98; Gaps 8;

OY 56 PLOFAMQKTSNYLAVTGADYIVKIFDRHQRSEINLPGNCVAMDMDKGVAVIAEK 115
DB 37 PIIRHW-RBNQHTVAVACANVTYIYDKKGNVLDALNPTGLIDIAMDEGDVLAIVAN 95
OY 116 SSCIYLMDANTKTKTQOLDNG---MRDQMSFLMSKVGSTLANGYKGNLXIYNHQTSTRI 172
DB 96 TQTLTMDVNSKNTDVTESGATSSKELPTCLAMSPSTPLVIGNNAGNIVNHTSRRI 155
OY 173 PVLGKHTKRTGCGNAENMLXALGDEKMITVSNQGDITROTQVRSXPNNQFLMKMD 232
DB 156 AMGKHQRQSVQITVTPBEDYVISCSDNLTSLVTLLEGTVSTTTNGEPTNNDY--GSVN 213
OY 233 DRTSAESMISVYLKKTFLFNLNLEPPNADLEPQDGNATVCGNMGDGKIMGFSCG 292
DB 214 GGGGQVTVWVSIVKIKIMLAHYNALDEPVNLQOEKGNHSHRWFDYILLGFPRG 273
OY 293 HFVVIHSTHGEIGELFOQRNKHNDLTSIAVSQTLNKVATCGDNCIKIQLDVKDMYVI 352
DB 274 YIISISAHNNEIGSELVSFLERGLASIAVSTSEKLLITIDNMVKKVDELITVTML 333
OY 353 LNLDEKNGKGLTSMWDDQLALSTORGLHVFLLKPLILDGASTRIAYLTSLEVTY 412
DB 334 TELERE-KNLSEIETVEDQILVAVSSQSGVLISFYTKMPTLAASVNSICVLTNLQTV 392
OY 413 ANPVGEPLTVSVDEPNEFVAVGLYHLAVGNNNRWY-----VLGNNAV 458
DB 393 VAEEVKKSGSTLELIEPTVMGLPLNLAIVANNNTVFEDYHTPQAQMAAQLOSTQSA 452
OY 459 KK-----LKDMYLGLTVASICLHSDYAALFEGKYQLHIESEIIDAQERETRLFP 510
DB 453 EKPITVAEPINRVELYLSVTNLIQNLVMAAANFSGRLRLHRIIN-----SDNVSIIEP 507
OY 511 AVDDCRILCHALTSDFLITGDTGVQVGYFYEDMQFVNDYRHPVSVKIIFPDPAQTRLV 570
DB 508 EARNRATLXYALTENFL----- 526
OY 571 FIDKSDGFCVPCVNDATVETLPDESPTIKGLVEMWPMKGVFIAYDDKVTYVYFHKDT 630
DB 527 -----FTTSYKGAIMEFTTIDKKNFAVFDSONIVFLLSKOH 563
OY 631 IGGAVIILAGSTKVPFAKHPILLYNGELTLCQOSQKVVNNIYSTGIFLSNLKDXPDEL 690

Db	564	IQGSGSVIYVATRLPHAVYVPLSLNKGIIVCMNSNGKLSLLSDSHKTSVSIIDSKSETVID	622
Qy	691	PMLANHMLTKR 701	
Db	624	DILTRSLIMHR 634	
RESULT	6		
Q9M097		PRELIMINARY;	PRT; 1205 AA.
ID	Q9M097		
AC	Q9M097;		
DT	01-MAY-2000 (TREMBLrel. 13, Created)		
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)		
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)		
DE	CG3069 PROTEIN (LD29485P).		
CN			
OS	Drosophila melanogaster (Fruit fly).		
OC	Eukaryota; Metazoa; Anthropoda; Tracheata; Hexapoda; Insecta;		
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;		
OC	Ephydroidea; Drosophilidae; Drosophila.		
OX	NCBI_TaxID=7227;		
RN			
RP	SEQUENCE FROM N.A.		
RC	STRAIN=BERKELEY;		
RX	MEDLINE=20196006; PubMed=10731132;		
RA	Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,		
RA	Amaratilake P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,		
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,		
RA	Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,		
RA	Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,		
RA	Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,		
RA	Abblitt J.F., Agbayani A., An H.-J., Andrews-Pfankoch C., Baldwin D.,		
RA	Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,		
RA	Beeson K.Y., Benos P.V., Berman B.P., Bhargava D., Bolshakov S.,		
RA	Borkova D., Borchen M.R., Bouck J., Brokstein P., Brothier P.,		
RA	Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,		
RA	Cherry J.M., Caley S., Dahlke C., Davenport L.B., Davies P.,		
RA	de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,		
RA	Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,		
RA	Durbin K.J., Evangelista C.C., Ferraz C., Fertlera S., Fleischmann W.,		
RA	Floeder C., Gabrielian A.E., Gang N.S., Gelbart W.M., Glasser K.,		
RA	Goelck A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,		
RA	Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,		
RA	Hoslin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,		
RA	Jatall B., Kalush F., Kapen G.H., Ke Z., Kennison A.G., Ketchum K.A.,		
RA	Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,		
RA	Laeko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,		
RA	Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,		
RA	Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,		
RA	Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,		
RA	Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclio J.M.,		
RA	Palazzo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,		
RA	Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,		
RA	Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,		
RA	Spryler E., Spradling A.C., Stapleton M., Strong R., Sun E.,		
RA	Sylvester R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,		
RA	Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,		
RA	Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,		
RA	Ye Y., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,		
RA	Zhang X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,		
RA	Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;		
RT	"The genome sequence of Drosophila melanogaster.";		
RL	Science 287:2185-2195(2000).		
RM			
RP	SEQUENCE FROM N.A.		
RC	STRAIN=Y, CN BW SP;		
RA	Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,		
RA	Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,		
RA	Gonzalez M., Guarlin H., Li P., Liao G., Miranda A., Mungall C.J.,		
RA	Nunoo J., Paclio J., Pargacs V., Park S., Phouanavong S., Wan K.,		
RA	Yu C., Lewis S.E., Rubin G.M., Celniker S.,		
CC	Submitted (OCT-2001) to the EMBL/Genbank/DBD databases.		
CC	-1- SIMILARITY: CONTAINS 2 WD REPEATS (TRP-ASP DOMAINS).		

QY	DB	Match	Score	Length	DB 5	DB 10	DB 20	DB 40	DB 80	DB 160	DB 320	DB 640	DB 1280	DB 2560	DB 5120	DB 10240	DB 20480	DB 40960	DB 81920	DB 163840	DB 327680	DB 655360	DB 1310720	DB 2621440	DB 5242880	DB 10485760	DB 20971520	DB 41943040	DB 83886080	DB 167772160	DB 335544320	DB 671088640	DB 1342177280	DB 2684354560	DB 5368709120	DB 10737418240	DB 21474836480	DB 42949672960	DB 85899345920	DB 171798691840	DB 343597383680	DB 687194767360	DB 1374389534720	DB 2748779069440	DB 5497558138880	DB 10995116277760	DB 21990232555520	DB 43980465111040	DB 87960930222080	DB 175921860444160	DB 351843720888320	DB 703687441776640	DB 1407374883553280	DB 2814749767106560	DB 5629499534213120	DB 11258999068426240	DB 22517998136852480	DB 45035996273704960	DB 90071992547409920	DB 180143985094819840	DB 360287970189639680	DB 720575940379279360	DB 1441151880758558720	DB 2882303761517117440	DB 5764607523034234880	DB 11529215046068469760	DB 23058430092136938400	DB 46116860184273876800	DB 92233720368547753600	DB 184467440737095507200	DB 368934881474191014400	DB 737869762948382028800	DB 1475739525896764057600	DB 2951479051793528115200	DB 5902958103587056230400	DB 11805916207174112460800	DB 23611832414348224921600	DB 47223664828696449843200	DB 94447329657392899686400	DB 188894659314785799372800	DB 377789318629571598745600	DB 755578637259143197491200	DB 1511157274518286394982400	DB 3022314549036572789964800	DB 6044629098073145579929600	DB 12089258196146291159859200	DB 241785163922925823197196800	DB 483570327845851646394393600	DB 967140655691703292788787200	DB 1934281311383406585577574400	DB 3868562622766813171155148800	DB 7737125245533626342310297600	DB 15474250491067252684620595200	DB 30948500982134505369241190400	DB 61897001964269010738482380800	DB 1237940039285380214769647769600	DB 2475880078570760429539295539200	DB 4951760157141520859078591078400	DB 9903520314283041718157182156800	DB 19807040628566083363314364313600	DB 39614081257132166726628728627200	DB 79228162514264333453257457254400	DB 158456325028528666906514914518400	DB 316912650057057333813029829036800	DB 633825300114114667626059658073600	DB 1267650600228229335252119316147200	DB 2535301200456458670504238632294400	DB 5070602400912917341008477264588800	DB 10141204801825834682008954529177600	DB 20282409603651669364017909058355200	DB 40564819207303338728035818116710400	DB 81129638414606677456071636233420800	DB 162259276829213354912132672468441600	DB 324518553658426709824265344936883200	DB 649037107316853419648530689873766400	DB 1298074214637066839297061379747532800	DB 2596148429274133678594122759495065600	DB 5192296858548267357118245518990131200	DB 10384593717096534714236491037980262400	DB 20769187434193069428472982075960524800	DB 41538374868386138856955964151921049600	DB 83076749736772277713911932303842099200	DB 166153499473544555427823866007684198400	DB 33230699894708911084556773201536839372800	DB 66461399789417822169113546402673678745600	DB 13292279957883564433822709280534751748800	DB 265845599157671288676454185610695035974400	DB 531691198315342577352908371221390071958400	DB 106338239663068515470581674242378191577600	DB 2126764793261370309411733484847583831555200	DB 4253529586522740618823466869695167166102400	DB 8507059173045481237646937339390334332243200	DB 1701411834609096247529387467878066866444800	DB 3402823669218192495058774937756133728889600	DB 6805647338436384990117579875512267577779200	DB 13611294676872769980235159511024535155554400	DB 272225893537455399604703190220490703111108800	DB 544451787074910799209406380440981406222217600	DB 108890357414982158841881260888196281244443200	DB 217780714829964317683772521776352562488886400	DB 43556142965992863536754504355270512477777280
----	----	-------	-------	--------	------	-------	-------	-------	-------	--------	--------	--------	---------	---------	---------	----------	----------	----------	----------	-----------	-----------	-----------	------------	------------	------------	-------------	-------------	-------------	-------------	--------------	--------------	--------------	---------------	---------------	---------------	----------------	----------------	----------------	----------------	-----------------	-----------------	-----------------	------------------	------------------	------------------	-------------------	-------------------	-------------------	-------------------	--------------------	--------------------	--------------------	---------------------	---------------------	---------------------	----------------------	----------------------	----------------------	----------------------	-----------------------	-----------------------	-----------------------	------------------------	------------------------	------------------------	-------------------------	-------------------------	-------------------------	-------------------------	--------------------------	--------------------------	--------------------------	---------------------------	---------------------------	---------------------------	----------------------------	----------------------------	----------------------------	----------------------------	-----------------------------	-----------------------------	-----------------------------	------------------------------	------------------------------	------------------------------	-------------------------------	--------------------------------	--------------------------------	--------------------------------	---------------------------------	---------------------------------	---------------------------------	----------------------------------	----------------------------------	----------------------------------	------------------------------------	------------------------------------	------------------------------------	------------------------------------	-------------------------------------	-------------------------------------	-------------------------------------	--------------------------------------	--------------------------------------	--------------------------------------	---------------------------------------	---------------------------------------	---------------------------------------	--	--	--	--	---	---	---	--	--	--	---	---	---	---	--	--	--	--	---	---	---	--	--	--	--	--	--	---	--	--	--	--	---

Db 764 YGEFEAEKLYDADRDRLAIELRMTLCDWFRVODLYRMGGSGVSDQOMELMREIGHNF 823
QY 810 APDQIPFISKEAIOLEFAGDYVNLALHYEKITGDNKEHDEACLAGVAMQSIKRMGDIR 869
Db 824 ANLRSMESAREYKESHYLEGYEALYHLE-----QDDLEK 860
QY 870 GYNQALNHRSVLKRDCGAILNKOFSEADQYLEKGLYDKAASV--YIRSKNNAKVG 927
Db 861 CV-ERLPEKSPLLPK-LAEMLASVGMSEAVQAHLR--FGDQKAAVATCVNLRQMGAEVE 916
QY 928 LLPHVSPPKIHQYAKAKE--ADGRYKAAVVAAYENAKMOSVIRIYLDHLNPKKAVNI 984
Db 917 LAQRFQLPQVOTLLAKHAAQLQEBRLEKEALIMORNA----- 953
QY 985 VRETSQDCAKAVAFRFLQIDGYSALQFLVMSKCNNEAFTLAQOONKMEIYADIISSED 1044
Db 954 ---GRHDAARLSQ--MAERQEKRAPLLRIKKLYVALALLAEHLK-----AVA 999
QY 1045 TTNEDYOSIALYEESEKRYLQAGKFFLLCGQYSRAIKHFLKCPSSSEDNVALMALETVOG 1104
Db 1000 TTEIDYAS-----GRNTLL-----DSIALE----- 1019
QY 1105 AKDELLTQQLDHLLEGENDMKDAKYFLRYMALKOYREAAQTALITAREQSAGNYRN 1164
Db 1020 -----DAALERLWHCAEAYH-----PMLLAQRLQRLFG----- 1047
QY 1165 AHVDLFMYAEKLSOKIRIPSEMATNLMLHSYLVKIHVKNQDMKGAIRLVANNIS 1224
Db 1048 -----IVHSAAVTVAVRLRYEDVLP----- 1067
QY 1225 KPPSHIVILSTVIECHNACIKNSAFSAMLRPREYRSKID---AKYKKKIGGMVRR 1280
Db 1068 --PEHYISLMLASCADRAFGTCSKAF-----MKLEODANLPEAYTLQRYEELAGIFAK 1119
QY 1281 PDISEIEATPPCPRCFKFLPCELCGCKNSIPYCTATGNNMK---DDW--TVCPRC 1335
Db 1120 YDEPTTGDRVDCYSGVPRVDPSSPCPCBNARFACISSGKPTIQTPTNINICTTCHNC 1179
QY 1336 DEPALYSEIKMLNTESTCPMC 1357
Db 1180 AAPT-----EISRHTCPLC 1194

RESULT 7
Q9JUK3 PRELIMINARY: PRT: 1749 AA.
AC Q9JUK3
ID 09JUK3
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE SELECTIVE LIM BINDING FACTOR.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20250882; PubMed=10788441;
RA Howard P.W., Maurer R.A.;
RT "Identification of a conserved protein that interacts with specific
RT LIM homeodomain transcription factors";
RL J. Biol. Chem. 275:13336-13342(2000).
CC -1- SIMILARITY: CONTAINS 5 WD REPEATS (TRP-ASP DOMAINS).
DR EMBL: AF226993; AAF68274.1; .
DR InterPro: IPR002106; AA_trna_ligase-II.
DR InterPro: IPR001680; WD40.
DR Pfam: PF00400; WD40; 5.
DR SMART: SM00320; WD40; 4.
DR PROSITE: PS00339; AA_trna_ligase-II_2; UNKNOWN_1.
KW Repeat; WD repeat.
SQ SEQUENCE 1749 AA: 197601 MW: B72C79207F5C0A5 CRC64;

Query Match 4.0%; Score 289; DB 11; Length 1749;
Best Local Similarity 19.1%; Pred. No. 1.2e-10;
Matches 279; Conservative 195; Mismatches 480; Indels 504; Gaps 70;
QY 39 VEMKRIFSLLETWIGAPQ-FANQKTSQNYLAIVGADYIVAFIRHCKQSEINL-PGN 96
Db 1 MQLKHLRTLLSPQDAAKATCTMAWSQNNKAF-AVCTVDRVALLYDEHGRRRKFKSTPK 59
QY 97 -----CVAMDMDKDDVDLAVIAEKSSCIYLDAN-----TKTKSQLONGMD 138
Db 60 MKYGRKSYVKKMAKSPDSTKIA-IGQDNIITYYKIGEDMDKVKCKFTQ-----TS 113
QY 139 QNSFLMSKVGSFLAVGVYKGLXYNNQTSRKIPVLG--KHTRKPIYTCG----- 186
Db 114 AVTCLQW-PAEYIIVEGALGKVRLANFTKNSITYGDSYVALTNGSGKILSGHA 172
QY 187 -----WANEKALGDEKMITVSQEDDTIRQ 214
Db 173 DGTIVRYEFDDESGSESGKLVNHCOPYALAW-ATNSIVAAGCRRIVAYGKEGHVLTQ 231
QY 215 TQVRESEPNMOPFLKMDRDTSAESM--ISVVLG-KTLFPLNIN-----EPDNPADL 265
Db 232 FYSRDPQEREF-----TTAASPGQSYVLSYDRLRFPNNSPRSTIEKAPKEI 283
QY 266 EFQDQGNIVCYNMYGDG-RIMIGFSCGHFVVISHTGELGOEIFQA-----RNHKDNL 318
Db 284 ---ANLVTYVALAMKRDGSRCLAGTLCG-----GVEQPDCLRSIRYKKEEL 328
QY 319 TSIAVSQTLNKATAGDNCIKIQDLVDLKDNY---VILNLDENNGIG---TSLWTDG 371
Db 329 TYVGSQQLVKKLSSGTRVY-----LKSHTGYEV---BEVKILGERLYVAHTST 376
QY 372 QLLA-LSTQSGSLHFLTKPLPLDGACSTRIVALTSLLEVTVANVEGELPTVSVDVER 430
Db 377 LLLGLDLNTR-----LSLEIPWQSGGNGEKYFF-----ENK----- 406
QY 431 NFVAVGLYHLAGMNNRMAFVYLGNAVAKKLDMEYLGTVASICLHSDYAAALFEGKVL 490
Db 407 -----NVCMIENAGELTVEYGSNLSGVRTEFPNMP----- 438
QY 491 HLIESEILDQAE---ERETRLPRAVDCKRICLHALTSDFLYGDTGVQVYFIEDNQF 547
Db 439 HLISVIRNECORGMEDKKLAVLDIKTIAVD-----LIGGINIGITISHESRYDME 492
QY 548 VNDYRHPVSVKKIFPDPRGTRLVLFIDEKSDGFVYCPVNDATYEIPDFSPY--IKGLVME 605
Db 493 LNE-----TGHKLLFDRK-----LRLHLVDIESCSKTMILNFCSTVQ 530
QY 606 WPMKGVFLAYD-----DDKYTYVYFHKDTI---QGAK---VILAGSTKYPF 646
Db 531 WPGSDVLAQNRNSLCWYNIAPERYTMSIRGDVGLERGSGKTEVMYEGVTTVAY 590
QY 647 AHKPLLNLNGELTQGTQSGKVNIIYLSHGFSLNKDKGPDRLRMLAHNLKRFSDAM 706
Db 591 T-----LDEGLIEFTALDDGN--YTRATAFLETL-----EMTP----- 622
QY 707 EWCRIILNDEAANMLARACLHHEVEFAIRYVRIGNVGVLSBOIKI----- 756
Db 623 -----ETFAAMKTLTALAEARQLHTAERCSALGHVAKAFLEHETNDAQVREYGG 676
QY 737 --EDYNLAGHLAMFTNDYNLAQDLYLASSCPITALLEMRDLQHMDSALQKLHAPDI 814
Db 677 ECTDYYQVRAFLAMEKYNKLAEMIFLQONAVEAMDYOEHLRWDECIJAVEAGHPAL 736
QY 815 PLSKEVAYIOL-----EFAGDYVALAHYEKGITGDKKHEDEACLAGVAMS 861
Db 737 EKLIRDYIQRLLMDIQOERAGELQESQDGLAIVLYLK-----ALPKA 782
QY 862 IRMGDIRGVNQALNHRSVLKRDCGAILNKKOFSEAAQYLEKGLYDKAASVYIRSKN 921
Db 783 ARL-----VLTRE--ELLANLEVEHITTLALIKELYERA----- 815

OY	922	MAKGDLLPHVSPKRIHQ-----YAKKE-ADGKYKAAVAVAEAKMOSVITIID	973
OY	922	MAKGDLLPHVSPKRIHQ-----YAKKE-ADGKYKAAVAVAEAKMOSVITIID	973
Db	816	-----GDLEKIRNPORALEYCSCKNAFMAVELARLAEFVEVRLAEA--WG-----D	862
OY	974	HLNPEKAVNIVREPOSIDGAKMVAFFLOLGDYGSALIOFLVMSKCNFAF--TLAQOHN	103
Db	863	HL-----VQOKO-LD-----AIIINHYEARCSIKALAEALGAROM	896
OY	1032	KMEIYADIIGSDITTEDYOSIA-----LYEGERK-----RYLQAGKPEL	107
Db	897	KKAIYIIDLORNRASKYKYPFVAAOHVASLOEYELAEELTYRKGRDTRDAIDMYQAGR---	953
OY	1072	LCGOYRALKHFELCPSSDDVNALEMALEIVYGAKDELLNOLLHLGSDGMPDARY	113
Db	954	-----WEGAHKRLAMCKMPED-----VSPLYTQAQE-----MERQGY	987
OY	1132	-----LEFLY-----MALKQYRE-----AAQTALIAAREEOSACNYRN	116
Db	988	REAEELTYTVEBPDLATITMFKRHLVYDMDIRVQKHHPDLSLTHLGLKELEBRLQE	104
OY	1155	AHDVLEFSWAEELKQKIKIPSEW-ATNLMILHSYIIVKIHVKNGDHMK-----GAR	121
Db	1048	AE-----YHLEAQEWKATVNMVRSGLMEAEAVRVAKAGCANAHKHVAYLWAKSISGE	110
OY	1215	MLIRVANNISKFPFHYPILSTYIECHRRGLKNSASFAMLMREPYRSK---IDAKTK	127
Db	1102	AAVRLNLNLG-----LLEAID-HAA--DMCSEFAPFELSRLALKHRTPEIHLRYA	114
OY	1272	KKIEGMVRPRDISEIEEA	1289
Db	1150	MYLE-----DEGKFEEA	1161
RESULT	8		
O9HBG6			
ID	O9HBG6	PRELIMINARY;	PRT; 1242 AA.
AC	O9HBG6:		
DT	01-MAR-2001 (TREMBlrel. 16, Created)		
DT	01-MAR-2001 (TREMBlrel. 16, Last sequence update)		
DT	01-DEC-2001 (TREMBlrel. 19, Last annotation update)		
DE	WDR10P.		
GN	WDR10.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
OX	NCBI_TaxID=9606;		
RR	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=21134653; PubMed=11242542;		
RA	Gross C., De Baere E., Lo A., Chang W., Messiaen L.;		
RT	"Cloning and characterization of human WDR10, a novel gene located at		
RT	3q21 encoding a WD-repeat protein that is highly expressed in		
RT	plutlary and testis.";		
RL	DNA Cell Biol. 20:41-52(2001).		
CC	-1- SIMILARITY: CONTAINS 6 WD REPEATS (TRP-ASP DOMAINS).		
DR	EMBL, AF244930; AAC15427.1; -.		
DR	InterPro; IPR001680; WD40.		
DR	Pfam; PF004400; WD40; 6.		
DR	SMART; SM00320; WD40; 7.		
DR	PROSITE; PS50082; WD_REPEATS_2; 1.		
DR	PROSITE; PS50294; WD_REPEATS_REGION. 1.		
KW	Repeat; WD repeat.		
SO	SEQUENCE 1242 AA; 141968 MW; 40CIDE7BF34D5E8 CRC64;		

Db 966 FSVHRETF-----NISRFHLSPKDTFSGISKVKRILFTLAKQSKALGAYRLAHAYD 1021
QY 1264 S-----KIDPKYKKIE---GMRRPDISEIEATPTCPCKELLPECCCELLCPCCKRSITRY 1316
Db 1022 KLRGLYTPAFQKISLGLTLTTAKRPFHSE-----LVPLC-----Y 1059
QY 1317 CIATGRHMLKDMWTCPHCFPALYS-----ELKIMLNTSTGCP 1355
Db 1060 RGSINPLNINIGNCINCRQPFISASDYVLHVEFYLESGITDEAISIULDELVRP 1119
QY 1356 MCSERLAAQOLKKSIDCTOYLRTFEE 1380
Db 1120 KRDR-----OLETANNSSQILRLVE 1140

RESULT 9

Q22830 PRELIMINARY; PRT; 1758 AA.
AC Q22830;
DT 01-NOV-1996 (TREMBLREL. 01, Created)
DT 01-NOV-1996 (TREMBLREL. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLREL. 19, Last annotation update)
DE T27B1.1 PROTEIN.
GN T27B1.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_Taxid=6239;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Almscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Cratton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Shownkeen R.,
RA Smaildon N., Smith A., Sonhammer E., Staden R., Sulston J.,
RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA Watson A., Weinstock L., Wilkinson-Spoat J., Woldman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
elegans.";
RL Nature 368:32-38(1994).
RN [2]
RP SEQUENCE FROM N.A.
RA Wu X., Le T.;
RT "The sequence of C. elegans cosmid T27B1.";
RN Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
RL [3]
RP SEQUENCE FROM N.A.
RA Waterston R.;
RT Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
CC -1. SIMILARITY: CONTAINS 4 WD REPEATS (TRP-ASP DOMAINS).
DR EMBL: U41020; AA02332.1; -.
DR InterPro: IPR001440; TPR.
DR InterPro: IPR001680; WD40.
DR Pfam: PF00515; TPR; 2.
DR Pfam: PF00400; WD40; 4.
DR SMART: SM00320; WD40; 3.
DR Repeat: WD repeat.
KW SEQUENCE 1758 AA; 198406 MW; 894255C3FEF7B0A6 CRC64;
SQ

Query Match

Best Local Similarity 3.7%; Score 270.5; DB 5; Length 1758;
Matches 270; Conservative 220; Mismatches 493; Indels 463; Gaps 68;

QY 65 SCGYLAVTADYIVKIFDRHGRSEINLPGNCVADMDKDGVLAVIAEKSSCIYLMDA 124
Db 27 NSGRAIAACSDRSVALDENGVOKDRF-----TCKPID-----AKYGRKSTTVLC--- 71

QY 125 NTRKTSQDLNCRHDSQSFLLMSKVGSLFVATGVKGLXLYNHQTS--RKIPVLGRHKK-- 180
Db 72 -----TFSPDSSRIALIGOSDNVLFYKVGTSWNEKKVLYNKFVQPS 113
QY 181 RITCGCMANENLXALGGEKMTTVSNQEGDTTRQOVYRSEPNKMOFFLMKMDRISAAFS 240
Db 114 AVTCLSWPDDKILVGOQLDGKVRIG-----LIKTKNCSS-----LYKTD-----ET 154
QY 241 MISVVLGKRTLEFLNINEDPNADLEFOODFENIYCYNNYGGRIKIGFSC----- 291
Db 155 VSIOTHPKRRTSFVSAH-----OD-GSIIILYNNSSKTSQSTI--CLQVPPYVL 199
QY 292 ---GHEVVIS-----THGTGEGEITFOARNHKNLTSIASQTLKV--ATCG--- 334
Db 200 VFTNGLVATSDRRLVLYSTENGVOQCFDYNDQSEKSESSISCPDAQNVGVTCGVF 259
QY 335 ---DNCKR-----IDQLVD-----LKDQVY-- 351
Db 260 SVDCCLRRGMLKSRRETTYVAPSHVILRDVTDRTNVTISNGLAIDELKMGKDRYVIG 319
QY 352 -----ILNDEENKGLGLTSMDDGQLALSTORGLHVFLEKPLIGDACSTRILAYLTS 406
Db 320 YTSSTIITADTESQRSSELM-----QSGHEKFTFD---FNCC-----LIIN 360
QY 407 LLEVTVAN--PVEGELPITVSVDPENFVAVL---YHLAVGNMNRAMFYVIGENAVKRLK 462
Db 361 ABEVTVVEYGVDSIGWVTELTSPHLSVQLKAFKSKYGRNFDPS-----ESSVSYSE 414
QY 463 DMEYLGTVASICLHSDYAAALFEGKVQVHLISEILDAQEERTLRFPAVDKCRF---- 518
Db 415 HSDL--NTGQICVFSDFL-----QRTTSVFETIKAKKAFINDL 451
QY 519 -LCHALTSFLIYGTDTGVQVYFIEDMOFVNDYRHPVSVKRIFFDPNGTRLVFIDEKSD 577
Db 452 QLPVNAIPNLINQGOESINHTGAVDWIELNE-----RASKLYLNDKRS- 496
QY 578 GRYVCPVNDATYEIPDFSTIGVL-----WENPMKDKVFAYIYDDKYYTVFHKD--- 629
Db 497 -----KVLVDISSDQSFVLSFCTYVQVPMSPDVIAQSGDNLSIN--YNDLPE 545
QY 630 -----TIQ-AKVILIAGSKVFAHKPRLLYNGELTQOSQGV-----NIIYLSHGFLS 679
Db 546 QVTNKKIKGEVNAVLDAVRT-----EYVQEPYAKVAYELDNTOIE---FGA 590
QY 680 NLKDGPEDELRLPMLAHNMLKRFSDAMENCRIINDEAANNELARALHMEVEFARIYVR 739
Db 591 ALEKRDPRAVAFLESN--TSGTDAYSM-----WIRVAEMALEHGNLFVAQRCYA 638
QY 740 RIGNGVIVSLQIKGIED-----YNLAGHLMFTNDYINLAQDIYLIASSCPI 787
Db 639 AINDAKVYKRLDILEIADIASISIGDGTFFYKVRAMLAINGRKRKEAERIFLRQNDTE 698
QY 788 AALEMRDLOHMSALQJLAKHLAPDQIPISKEY-----AIOLEFA-GDYVNA 834
Db 699 SAIGWYTSLHKWDEALELKVLYNPEYEDLKTSLRALSDTQDSKAALAKLSDGDTLSA 758
QY 835 LAHYEK-----GITGDNKE-----HDEACLAGVAQMSIRMGDIRQVQALHPSRVILKR 884
Db 759 IDLYIKSNKPLSALSAANDSVLSQDENILRQIADSLV-----SQLYD 802
QY 885 DCGAILLENKQFSEAAQLEYEKLYIDKA-----ASVYIRSKNMKAVGDLRLHYVSSP 935
Db 803 KAGDVEYEKLKDFDKAVEYKGDGDAYGKAQIARFAPEKRVVTLLEQEWG----- 850
QY 936 KIHLOYAKAKEDGRY-----KEAVVAYENAKQWOSVIRIYLDHLNPEKAVNIVRET 988
Db 851 -LHLEYIGQYDAVNHVFNVDANLKKAVEAIPAKER-----PALSLIV--- 892
QY 989 OSIDGAKM-----VAREFLOLGDYGSALQFLYMSKCNNEAFTLAQOHNK---MEIYA 1037
Db 893 ENIQDQKVRTGYGGEIADHYSNKGDFERAERLFEVGEGLFNDAIMYGNKKNKIDAFRLSE 952
QY 1038 DIIGSEDTTNEDYOSTALYFEGEKRYLQAGKFFLLCGOYSRLAKLHFLKCPSSSEDNVAIEM 1097

```

Db 953 EFHGREATIS-SYLAKEADEDEHGRFAEABOLYTTIGMPHKAIOW----- 997
QY 1098 AIEYVOAKELLTNLOIDHLLGENDGMPKDAYLFRLYALKOYREAAOTAIILAREQ 1157
Db 998 --DRVGDDDLV---RLVERXGHEH--MHETRRF-----ATQYBERGD--LKAAREQ 1041
QY 1158 --SAGVNRNADVLFSMY--AEKLSOKIKI-PSEMATNL--MILHSY-----ILVKIH 1203
Db 1042 FLKAGDFRSA-----VMNYKSEMSDAVRIAKTEGGENMKOYLFWMKASIGDDAAVKLL 1097
QY 1204 VKRGDHHKQ-----ARMLIRVANNISKPPSHVPIILTSTVIECHRAGLKNSAF 1251
Db 1098 NKHGMLMEGIDFACETGAFDLAFDLARIG---AKDRMGTVHVRILATQLE-EEGRLEDAK 1153
QY 1252 SFPAAMLMRPY-----RSKIDAYKKKIEGMVRPRDISELEATPTCPFC 1296
Db 1154 HYVEATIKLNTYNTITWCOAVPSRPDLQRLGNRPDLAVEMFIRDNDWADERVAK--DHC 1211
QY 1297 KFLPE 1302
Db 1212 ESLLPD 1217

RESULT 10
ID 09HBG5 PRELIMINARY: PRT: 1292 AA.
AC 09HBG5:
DT 01-MAR-2001 (Tremblrel, 16, Created)
DT 01-MAR-2001 (Tremblrel, 16, last sequence update)
DE 01-DEC-2001 (Tremblrel, 19, last annotation update)
DR MDR10P-L.
GN MDR10.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21134653; PubMed=11242542;
RA Gross C., De Baere E., Lo A., Chang W., Messiaen L.;
RT Cloning and characterization of human MDR10, a novel gene located at
RT 3q21 encoding a WD-repeat protein that is highly expressed in
RT pituitary and testis.;
RL DNA Cell Biol. 20:41-52(2001).
CC 1- SIMILARTY: CONTAINS 6 WD REPEATS (TRP-ASP DOMAINS).
DR EMBL: AF244931; AAG15428.1; -.
DR InterPro: IPR001680; WD40.
DR Pfam: PF00400; WD40; 6.
DR SMART: SM00320; WD40; 7.
DR PROSITE: PS50082; WD_REPEATS_2; 1.
DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
KW Repeat; WD Repeat.
SQ SEQUENCE 1292 AA; 147385 MW; 85F6493ED3B4A496 CRC64;

```

Query Match 3.7%; Score 268.5; DB 4; Length 1292;
 Best Local Similarity 18.1%; Pred. No. 1.9e-09;
 Matches 250; Conservative 203; Mismatches 512; Indels 413; Gaps 57;

```

QY 92 NLRGNCVAMDMKDGVLAVIA-----EKSSC---TYLMDANTNKTSQLDNGMRDMS 141
Db 137 NQPG-----QKHNDALQCVSNPIRTHQLASCSDFGLMSPPQKSVSK--HKSSSKII 187
QY 142 FLMSKGVSLAVGTAGNLXIYHOTSRIPIV--LCKHTRKRTTCGW-----N 188
Db 188 CCGSTJNGOYALGCMFNGIISIRKNKEEYKTERPGSLSPIWISICWNPSSRMSESPWM 247
QY 189 AENLXALGGEDEMT-----TYSNDEGDTIRQTVRESEPNMOPFLKMDRFT 235
Db 248 RENDDA---EDVIYVNRIOETPLSTLKSAYVSSQSEAEDEEPEDEDSPR-----DDNL 298
QY 236 SAESMISVY-LGKKTLLF-LNLNPDNPADLEFQDQFGNIVCYNMVGDGR-IMIGFSCG 292

```

```

Db 299 EERNDLIAVADMOQKVSFYQSLGQIGKDRALNFDP-----CCISYFTGGEYILG---G 350
QY 293 HEVAVISTHNGELGOELFOARNHNDNLSTAVSOTLKNVATCGNCKIKIDVLKRMVYA 352
Db 351 SDKQVSLFT-----KDGV-----RLGTVG----- 369
QY 353 LNDENKGLGTLSTWDDGOLLASTORGSLHVFILKPLTGACSTRILAYLSTLEVTY 412
Db 370 ----EONSWMVTOQAPRDSYVYVGGODGISITFYQLIFSVHLYDRARIRSMYDVLY 425
QY 413 ANPVEGELPTYSVDPEPNEVAVGLYHLAVGMNRRAMFYVLGENAVKLLKMEYLGTVAS 472
Db 426 QHLIT-EOKYRIKCEKLVKIKAIYRRLAIQLPEKILIELYS---EDLSMHY----- 475
QY 473 ICHSDYMAALFEGKQVQHLIESEILDQEEER-ETRLFRAYVDK-----CRILCH 521
Db 476 --RVKERRIKKFECNL-LVVCANHILCOEKRLOQCSFGSVKBEROMESLIRYIKVIG 532
QY 522 ALTSDFLIYCTDGTGVVQYFYIEDMQFVNDYRHRVSVKIPPDNGTRLYFIDEKSDGFY 581
Db 533 PRREGILVGLKNGQILKIFVDNLPAIVLLKQATVAVRCLDMSARKKLAVDENDCLVY 592
QY 582 CPVNDATYELIPDFSPTIKGYLVEMNMPDKGVT--AYDDKVVYVYVHKDPTIGARVILA 639
Db 593 ---DIDTKELLQEPNANSVAMNTQCEDMLCFSGGYLNKASTFPVHROKLOG---FVY 646
QY 640 GSTKVPFPAHKLPLLYNG-ELTCOTQSGKVNNTLSTHGFSLNKDKQPDRLRMLAHNL 698
Db 647 G-----YNGSKIFC-----LHV--SISAVEPDSAPMYQY-LD 677
QY 699 LKRFSDAMEWCRIILNDEAANELARACLHMEVEFATRYVRRIGN---VGIVSLDQIK- 754
Db 678 KILFKRAYQIACGVYDTRMRELALEBGLDETAKKATIRQDLRYLELISIERKK 757
QY 755 -GIEDYLLAGHLAMFTNDYNYLAQDLYLASSCPIALLEMRDLQHDLSALQAKHLAPQ 813
Db 738 QGFTNNDLFLADVFSYQGFKEHAALKYRSGHENLALEMVTDL----- 780
QY 814 IPIFSKEVATQLEFADGVYVALHKEGTTGDMKKEHDEACLAOVAOMSTRMGDIRGVNG 873
Db 781 -----CMFYVADFLG-----SGDKE-----TKMLITQADMAR--- 810
QY 874 ALKHPKVLKRDGALLIENKQFSEAAQVLYEKLGYDKASVYIRSKNMAKVGDLLPHVS 933
Db 811 -----NKEPKAAVEMTISGHNKALEI-CGDHGV---DMLIDI- 847
QY 934 SPKIHLOYAKAKEDGRYKAAVVAENAKOMQSVIRIYDLNNEPEKAVNIVRETSLDG 993
Db 848 -----ARKLDKREPRLL-----LCATYLLKLLDSPGYA----- 875
QY 994 AKVAVARFLOLDGYSALQPLVMSKCNNEAFTLAQGNKKKEIYADIIGSDTTNEDYQSI 1053
Db 876 ----AETYLKMGDLKSLVQHLHVEFQRMDEAFALGEKHPERK-----DDITMPY 919
QY 1054 ALYFEGEKRYLOAGKFFLLCGQYSRALKHLKCPSSSEDNVAIEMAIETVGOAKDELLTNO 1113
Db 920 AQMLAENDREEDQAKFHKKRGQREAV-----QVLEQLLNNA 956
QY 1114 LIDHLGENDGMPKDAYLFRLYALKOYREAAOTAIILAREQSAGNYRNADVLFSMY 1173
Db 957 VVESRF-----NDAAVYV--WMLSMOCLDIAOP--AOKDTMLKLFYH-----FQRL 999
QY 1174 AELKSOKIKIRPSEMATNLMLHSYILVKKVKNDDHKKGAMLLIRVANNISKPPSHVPI 1233
Db 1000 AEL-----YHGTHAIHRTEDPFSVHRETTLE---NISRELLSLPK 1038
QY 1234 LSTVIECHR---AGLNSAFSPAAMLMRPYVS---KIDAKYKKKIE---GMVRRPDI 1283
Db 1039 DTPSGISIKVAILTLTKAQSAALGAVRLARHAYDKLRGLYIPARFGKSELGTILIRKPF 1098
QY 1284 SEIEEATTPCPFCFLLPECELLCPGCKNSIPYCIAATGRHMLKDDMTVCPHCPALYS- 1342

```

DB 1099 HUSEE-----LVPLC-----YRCSTNNPLNINIGVINCINCPFISSA 1136
QY 1343 -----ELKIMLNTSTCPMCSERLNAOLKKISDCTOYLRTEE 1380
DB 1137 SSYDVHLVEFYLEBSITDEBAISLIDLEVLRRKRD-----QLEIANSSQILRLVE 1190
RESULT 11
ID 09UF80 PRELIMINARY: PRT: 1198 AA.
AC 09UF80:
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE HYPOHETICAL.137.1 KDA PROTEIN (FRAGMENT).
GN DKFZP434K016.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxId=9606;
RN (1)
RP SEQUENCE FROM N.A.
RC TISSUE=TESTIS;
RA Koehler K., Beyer A., Mewes H.W., Gassenhuber J., Wiemann S.;
RL Submitted (DEC-1999) to the EMBL/Genbank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 5 WD REPEATS (TRP-ASP DOMAINS).
DR EMBL: AL133565; CAB63718.1; .
DR InterPro: IPR001680; WD40.
DR Pfam: PF00400; WD40; 5.
DR SMART: SM00320; WD40; 3.
DR PROSITE: PS50082; WD_REPEATS_2; 1.
DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
KW Hypothetical protein; Repeat; WD repeat.
FT NON_TER 1
SQ SEQUENCE 1198 AA; 137080 MW; EF3D17BB9B5DE88C CRC64;

Query Match 3.6%; Score 263; DB 4; Length 1198;
Best Local Similarity 18.2%; Pred. No. 4.1e-09;
Matches 245; Conservative 197; Mismatches 506; Indels 396; Gaps 54;

QY 116 SSCITVMDANTKTSOLDGMRDQMSFLMSKVSFLAVGTAKGNLXIYHQTSTRIY- 174
DB 70 SSDPGLMSPEQKSVSK--HKSSKIIICSWINDGOTLALGMENGIISLNKNGEEVKIE 127
QY 175 -LGKHTKRTTCGCW-----NAENLXALGGEKMI-----TVSNQCG 209
DB 128 RRGSLSPITWISICMNPSSRMSFMMNRNEDA--EDVIVNYIQEIPSTLSKAYSYSOG 184
QY 210 DTIRTOTVSEFXNNQOFLMKDDRTSAESMISVV-LGKTLFF-LNLNPDNPADLEF 267
DB 185 SEAESEEPEDDSR-----DDNLEERNDLAVADMCKVSKFYOLSGKQIGKDALNF 238
QY 266 QODFGIVYNNYGGGR-IMIGESGHEFVISTHNGELGQELFQARNHNDNLSTAVST 326
DB 238 DP-----CISTFTYGETYLLG--GSDKQVSLFT-----KQGV----- 269
QY 327 LNKVATCGDNCKIRIODVLVDKDMYVILNDEENKIGLSTWTDGQLALSTORGSLHVF 386
DB 270 --RLGTVG-----EQNSWWTQOAKPDSYVYVVGQDGIISFY 305
QY 387 LTKLILGDACSTRIVAYLTSLLEVIVANPVEGELPTTVSVDVPEPNFVAGLYHLAVGMNN 446
DB 306 QLIFTVHGLYKDRYAYRDSMTDVIYQHLIT-EOKYRIKCKELIVKIAIYRRRLAIQLPE 364
QY 447 RMFVYVLGNNAKXKLDKMYLGTVASICLHSDYAAALFEGKQVHLISELIDAEER-E 505
DB 365 KLILIELIS--EDLSDMY-----RVKKEIIKKFECNL-LVVCANHIIILCDEKRLQ 412
QY 506 TRLPFAVDK-----CRILCHALTSDFLIYGTDTGVGVQYFIEDMOPVNDYRHVF 555
DB 413 CLISFGVGEREMQMSLRIRYIVIGGPPREGILVGLKNGQILKIFVDLFAIVLLKQAT 472

QY 556 SVKTIFFPDNGTRLVEFIDEKSDGFYVCPYNADATYELIPDSFPTIKGVLMENMDKGVFI- 614
DB 473 AVRCLDMSASRKKLAVVENDTCLVY---DIDTKELLFQEPANSVAMTQCEDMLCSFG 529
QY 615 -AYDDDKVYTYFHKDTTIGAKVILIAGSTKVPFAKRPILLYNG-ELTCOTGSGKANNIYL 672
DB 530 GGYLNKASTFVPHROKLOG---FVVG-----YNGSKIFC----- 561
QY 673 STHGFLSNLKDGPDELRLAHNLMLKRFSDAMENCRILNDEAMNELARACLHMEVE 732
DB 562 -LHVF--SISAVEVQSAAMYQY-LDKILFKAYQIACIGVDTQMRRELAMALGLDPE 617
QY 733 FAIRYVRIRGN--YGIWVSLQIR--GIEDYNLAGHLAMFTNDYNLAQDLYLASSCPI 787
DB 618 TAKKAFIRVQDLRYELIISIERKKRGFTNNDLELADVFSYQGFHEAKYKRSCHEN 677
QY 788 AALENRDLQHHDSALQAKHLAPPOIPITSSEYALQLEFADYNNALAHYKGTGDNK 847
DB 678 LALEWYTDL-----CMFEYARDFLG-----SGDPK 702
QY 848 EHDEACLAGVAQMSIRMGDIRGVNQAALNHPRSVLKRDGALILEMKNQFSEAAQLEKGL 907
DB 703 E-----TKMLITRKQADMAR-----NIKEKKAVENTYISAG 732
QY 908 YDKAASVYIRSKNAKYGDDLPHVSSPKIHQYAKAKEADGRYEAAYAVENAKOMQSV 967
DB 733 EHVAKIEI-CGDHGV--DMLIDI-----ARKLDKAEERPEPL-----L 767
QY 968 IRIYIDLNNPKAVNIARETQSLDGAKMVAREFLOLGDYGAIDQFVMSKCNNEAFILA 1027
DB 768 CATYLRKLDSPGYA-----AETYLKMGKLSLVQJHETQWMDAFAIG 811
QY 1028 QOHNMEIYADIGSEDTTNEGYOIALYFESEKRYLAGKFFLLCGOYSRALKHFLKCP 1087
DB 812 EKHPERK-----DILYMPYAOQLAENDPFEBAQKAFHKAQRGREAV----- 852
QY 1088 SSEDVAIEMAIETYQAKDELTLNOLDHLLGENDGMPKAKYIFRLYMALKQYREAAQ 1147
DB 853 -----QVLEQLTNNNAVRESR-----NDAAVYV--WMLSMOCLIDIAQ 887
QY 1148 TAITIAREOSAGNRYNADVLFESYAVELKSQIKIRPSEMAVNLMLHSYLVKTHVNG 1207
DB 888 DP---AQKDTMLGKTYH-----FORLAEI-----YGHYAIHHTEDP 922
QY 1208 DHMKGARMLIRVANNISKEPPIVILITSTVIECHR---AGLKSASFSAAMLMPREYRS 1264
DB 923 FSVHREPELF-----NISRLHSLPKDPSPGISVKYILFTLAKQSKALGAYRLARHAYDK 978
QY 1265 ---KIDAKYKKKIE---GMVRPDISETIEATTCPPCKFLLPCCCELLCPCKKSTFYC 1317
DB 979 LRGLYIPARFQKISIELGLTTLIRAKPFHDEE-----LVPLC-----YR 1016
QY 1318 IATGRHMLKDMTVCPRHCPFPALYS-----ELKIMLNTSTCPM 1356
DB 1017 CSTNPNPLNINIGVINCINCPFISSASDYDLHVEFYLEBSITDEBAISLIDLEVLRRK 1076
QY 1357 CSERLNAOLKKISDCTOYLRTEE 1380
DB 1077 RDDR-----QLEIANSSQILRLVE 1096
RESULT 12
ID 09HAT9 PRELIMINARY: PRT: 1241 AA.
AC 09HAT9:
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE SPG PROTEIN.
GN SPG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Shan Y.X., Li J.M., Sha J.H.;
 RT "The research of spermatogenesis related genes."
 RL Submitted (Sep-2000) to the EMBL/GenBank/DBJ databases.
 CC - SIMILARITY: CONTAINS 6 WD REPEATS (TRP-ASP DOMAINS).
 DR EMBL: AF302154; MAGI3415.1;
 DR InterPro: IPR001680; WD40.
 DR Pfam: PF00400; WD40; 6.
 DR SMART: SM00320; WD40; 7.
 DR PROSITE: PS50082; WD_REPEATS_2; 1.
 DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
 KM Repeat; WD repeat.
 SO SEQUENCE 1241 AA; 141824 MW; 6C3C543369A6BDP5 CRC64;

Query Match 3.6%; Score 263; DB 4; Length 1241;
 Best Local Similarity 18.2%; Pred. No. 4.3e-09;
 Matches 245; Conservative 197; Mismatches 506; Indels 396; Gaps 54;

QY 116 SSCIIYMDATNKTSDIENGMRDMSFLMSKSGSLAVTGKGLXIIYHQTSPKIPV- 174
 DB 113 SSDFGLMSPEOKSVSK--HKSSSKIIICSWTNDGQYLALGFMGIIISIRKNGEEKYKIE 170
 QY 175 -LCKHTKRTICGCM-----NAENLXALGEDKMI-----TVSNQEG 209
 DB 171 RRGGSLSPISTICMNSPSSKESFMNRENEDA--EDVIYNTYQETIPSTLKSAYVSSOG 227
 QY 210 DTRQGVQSEPNMOMFLMKMDRTSAESMISV-LGKRTLEF-LNLNPNADLEF 267
 DB 228 SEAESEEPEDSDSPR-----DNLERNDDILAVADMCGKSPYVSGKQIGKDRALNF 281
 QY 268 QDQEGNVICVMYGDGR-IMIGFSCGHFVVISHTGELGEIQARHKNKNTLSIANST 326
 DB 282 DP-----CCISYFTKGYILLG--GSDKOVSLFT-----KGGV----- 312
 QY 327 LNKVATGDCMCIKIQDVLVDKDMYIILNDEENGLGTLSTWTDGQLLSTORGLHVE 386
 DB 313 --RLGTG-----EQNSWVTQCAKPRSNVYVVGCGCGGTISFY 348
 QY 387 LTKPLIGDACSRIAYLTSLEVTYANPVGELPITVSVDENPVAVGLYHLAVGMN 446
 DB 349 QLIFSTVHGILYKDRYAYRDSMTDVIYQHLIT-EDQVKRIKCELVKIAIYRNRLAIDLE 407
 QY 447 RAMVYVGENAVKVKLKMELGTVASTCLSDYAAALFEKGVOLHLESELDAQDER-E 505
 DB 408 KILYELYS--EDLSDMHY-----RVKEKIKKFECLN-LVGCANHIILCOEKRLQ 455
 QY 506 TRLEPAVDK-----CRILCHALTSDFLIYGTDTGVVQYFYEDMOFVNDYHRPV 555
 DB 456 CLSSSVKKEEMOMESLIRIKYIGVPRGREGVLGKNGOILKIFVDNLFAIYLKQAT 515
 QY 556 SVKKIFPDNGRLVFIIDEKSDGEVYCPVNDATYETPDSPITKGLVEMWMPDKGFI- 614
 DB 516 AVRLCLDSASRKKLAVVANDENDTCLVY--DIDTKELLFOEBNANSVAMNTOCEDMLCFSG 572
 QY 615 -AYDDKVVYVYVPHKDTIQCAKAVTLASTKVPFAHKPLLLYNG-ELTCQOSGKVNIIYL 672
 DB 573 GGYINIKASTFPVHRQKLG--FVVG-----YNGSKIFC----- 604
 QY 673 STHGFLNLKDXGDEPLMLANLMLKRFSDAMEMCRILNDELAAMMELARACHHMEVE 732
 DB 605 -LHVF--SISAVEPQASAPMYQ-LDRKRLKFAEQIACLGVTDDIMELAMEALEGLDFE 660
 QY 733 PAIRVYRIGN--VGIVMSLEQIK-GIEDYNLLAGHLAMFTNDYMLADLYLASSCPI 787
 DB 661 TAKKAFIRVQDLRYLELISIEERKKRGGETNNDFLADVESYQCKFHEAALYKRSCHEN 720
 QY 788 AALEMRRDLOHMSDALQAKHLADQIPFISKEVAIQLEFAGDVVNLAMHAKERTIGDNK 847
 DB 721 LALEMVYDL-----CMFEYAKDFLG-----SGDPK 745

QY 848 EHDEACIAGVAQMSIRMGDIRGVNQALKHPRSYLRKRDCAILEMKNKOFSEAAQLEYEKL 907
 DB 746 E-----TKMLITKQADMAR-----NIKEPAAYEMVTSAG 775
 QY 908 YNDKASVYIRSKNMAKVGILLPHVSSPKIHLQAKKEDGKRYEAVYENAKKQMSV 967
 DB 776 EHVKALEI-CGDHGW--DMLIDI-----ARKLDKKEREFLL-----L 810
 QY 968 IRIYDLNPNPEKAVNIVRTQSLDGKAVARPELOLGDGSAIOPLYMSKNNEAFTLA 1027
 DB 811 CATYLLKLDSSPYA-----AETYLKMGDLKSLVQLHETORMDENAFALG 854
 QY 1028 QOHNMKEIYADIIIGSEDTTNEDYOSIALYFEGEKRYLOAGKFFLLCGOYSRALKHFLKCP 1087
 DB 855 EKHPEFK-----DDIYMPYQWLAENDRFEQAARHKAGROREAV----- 895
 QY 1088 SSEDNVAIEAIEYTGQAKDELLTNQILDLGENDGMPYDAKYLFRLYALKOYREAQ 1147
 DB 896 -----QVLEQLTNNAVAESRF-----NDAAVYV--WMLSMQCLDIDIAQ 930
 QY 1148 TAITIAEBSAGYRRAHNVLFMSVAELKSQKIKIPSEMATNMLHSYILVKIHVKG 1207
 DB 931 DP--AOKDMLGFEYH-----FORLAEL-----YHGTHAIHRTEDP 965
 QY 1208 DHMKGARMLIRVANNISKPSHIVPILTSYIECHR--AGLKNASAFSFAAMLMPPEYRS 1264
 DB 966 FSVHREPTLE-----NISRFLLHSIPKPTPGISGVKILFLFLAOSKALGAVRLARHAYDK 1021
 QY 1265 ---KIDAKYKKKIE--GAVRRPDISEIEBATTPCFCKFLLPCECLCPGCKNSIPYC 1317
 DB 1022 LRGLYIPARFOKSTIEGLTLIRAKPFHDSSE-----LVPLC-----YR 1059
 QY 1318 IATGRHMLKDMWYGCPCDPEALYS-----ELKIMLMTESPCPM 1356
 DB 1060 CSTNPLNLNDGNCVCIKCRPFIIFSASSYVHLVFRYLEBGTDEPAISLIDLEVLPRK 1119
 QY 1357 CSERLNAQLKKSIDCTQYLRTEE 1380
 DB 1120 RDOR-----QLEIANNSQILRLIVE 1139

RESULT 13
 060332 PRELIMINARY; PRT; 1462 AA.
 AC 060332;
 DT 01-AUG-1998 (TrEMBLrel. 07, Created)
 DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE KIAA0590 PROTEIN.
 GN KIAA0590.
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Tissue=Brain;
 RX MEDLINE=98290545; PubMed=9628581;
 RA Nagase T., Ishikawa K., Miyajima N., Tanaka A., Kotani H., Nomura N.,
 RA Ohara O.;
 RT "Prediction of the coding sequences of unidentified human genes. IX.
 RT The complete sequences of 100 new cDNA clones from brain which can
 RT code for large proteins in vitro."
 RL DNA Res. 5:31-39(1998).
 DR EMBL: AB011162; BAA2516.1; -
 DR InterPro: IPR002885; PPR.
 DR SMART: SM00320; WD40; 1.
 DR PROSITE: PS50082; WD_REPEATS_2; 1.
 DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
 KM Repeat; WD repeat.
 SO SEQUENCE 1462 AA; 165196 MW; B05CF4325F364F0C CRC64;

Query Match 3.5%; Score 256.5; DB 4; Length 1462;
 Best Local Similarity 17.3%; Pred. No. 1.6e-08;
 Matches 247; Conservative 219; Mismatches 485; Indels 475; Gaps 64;

54 GAPIDPAMOKTSGNVLAV-----TGADYIVKIFDRHQ--KRSELNPGKNCVADMDK 104
 16 GSPFSLMHPVH--PFLAVAYISTTSGS--VDILEGECVPDIVERFERFVAAVLCMPH 71
 105 DGDVLAVIAEKSSCIYLPDANT-----NKTSQLDNGM----RDQMSFLMSKVSFLAVG 156
 72 TRVLAV-----GWEIYGEVTVENKODKEGHTMPRLHTADITVLRNPSNCLSGD 122
 157 VGNLKIYHNQSRKI---PVL---GKHTKRTGCGW-----NAENL-----XALGED 199
 123 RLGVLLMLRLDQGRVQGRPFLKHEYGKHLTH---CIFRLPPGEGDVLQAKAAVSGDE 178
 200 KMTVSNQEGDITROPVASEPANNQFELMKD-----DRISAAMISVVLGKTLTF 253
 179 KALDMNNKSSSGSLKMGSHGLLFVSLMDGVHYVDEKGTQVVSADSTIOMLFY 238
 254 LLINE-----PDNPAD---LEFOQDFGNIVCYMWGDMGIMIGSCG 292
 239 MEKRELVVYVTEMLRLSLIVTPREGKAEEVMKYLSGKTR-----RADIALIG 288
 293 HEVYISTHFG-----ELGQETFOARNHK-----DNLTSVAVSOTLNVKATCGDN--- 336
 289 SLIVMAVGEALFMDIERGENYILSPDEFGEKENMNCVCYCKVKGILLAAIGDRGRV 348
 337 --CIKIDLIY-----DLKDMYVILNIDENKKGITLSTWDDQGLLALSTQSGSLHFLRK 389
 349 AMRKVPDLFGSGAGCKRMALQPTLEOGNTTQIQMSRKMLLVANS-----VIS 400
 390 LPLIAD-ACSTRIAYVLSLEYVANPVEGELPIFVSVDPERFNAVGLYH----- 439
 401 VALISRASSHHQVAAQVNS-----PSLLNV---CFELSTGAHSLRTDMHS 447
 440 -----LAVGNMNRAMTYVJGNAVKLKDMEYLGVAASICLSHDYAAALFEGVQLH 491
 448 GFVATKDAVAVMNRQVAFELSGAARISAG--TFLCETPVLAMHBEENVYTESNVOVR 505
 492 LIESELDAQEEERETRLPFAVDDKCRILCHALTSIDLTIGTGVGYQVYIEEDMOWNDY 551
 506 TWQGYVKOLLSETSEIGNCFELDIC-----GNFLVGTDLALFKSDLSRR--AKAH 556
 552 RHPVSVYKIFP-----DENGTRLVFIDEKSDG-----FVYCPVNDATYEIPDFSP 596
 557 CCSRSLAELVPGVGIASLRCSGSGTISILPSKADNSPSDKICFYDVEDMDYTFVDF-- 614
 597 TTKGVLMEMPMDDKGVFIAYDDKVITYVFHKDTIGAKAVIAGSTKVPFAH-----KFL 651
 615 -----KTGOIDRRETLSEFNEOE--TNKSHLFEVDEGLK-----NYVUNHFWQOSEPR 659
 652 LLY-----NGELTCOTQSGK-----VNNIYLS--THGFLSLMKXGPD 687
 660 LVCVCAVQETPSQPOSANG-----QPQDGRAGPADVILISFFISEHGF--LHESFP- 712
 688 ELRPMLANLNM-----SDANEMKRIILNDEAAMNLAACLHMEVEPAIRYRIGAV 744
 713 --RPATSHSLDMEVYUYFTRKPEADREDEVEPCHHIIPQVNSRPLRDFGLJEDCDK 770
 703 -----SDANEMKRIILNDEAAMNLAACLHMEVEPAIRYRIGAV 744
 771 ATRDAMLHESFVTTIGDMEAFKSIKLISEAVMENMAMCYKTORLDAVA--KYC--LGMN 827
 745 GIVMSLEQKJIEDVNLNLAGHLAMFTNDYNLAODL--YLASSCPIALALEMRDQH----- 798
 828 GARAGARALREKQEPLEEARVAVLATQUGMLEDAEQLRKC-----KRHDLNLFYQA 881
 799 ---WDSALQALAKHLAPDQIPFTISKEAYALOLEFAGDVVNALAHYKGTGDNKEHDEACIA 855
 882 AGRMQEALQVAAHHDRVHLRSTYHRYAGHLSEASDCSRALSYEKSDDT----- 929

856 GVAQMSIRMGDIRGVNQAQKHPSRVLRKDCGAILIENMKOFSEAAQLYEKGLYIDKAASV 915
 930 -----HREFPV--RMLSEDLPSL-----ELYVK----- 951
 916 YIRSKN-----NAKVGDLPLRHVSSPKTHILOAKAKREADGRYKENVAYENAKOMOSIRY 971
 952 -MKDKTLMRWA-----OYL---ESQGMDAALHYELARHFSILVRH 991
 972 LDHLNPEKAVNIVREFQSLDCAKVAFFLOLGDGSAIOFLVMSKCNNEAFTLAAQHN 1031
 992 C-FQGVQQAQJANTGNGLASVYHARQYEOQEEVGAQVHHYTRQAKNRIRLCKEN- 1049
 1032 KMEIYADITIGSED-----TTNEDYOSIALYFEGEKRYLOAKGFELL---CGQTSRAL 1080
 1050 -----GLDDQMLNALLLSSPEDEIMIVARY--EEKGVOMDRVAVMLYHRAHGFSKAL 1098
 1081 KHEFLKPSSEDNVAIEMALEIYGAQKDELITNOLDHLLGENDMPKDAK--LRLVYA 1138
 1099 ELAF---ATQGFVALQDLIEDDETSDPALLARCSDFI-----EHSQYRAVELLIA 1148
 1139 LKQYREAQTAIIRAEOSAGNYRNAHDVLFMSYAELEKSORIKIPSEMATIMILHS-- 1196
 1149 ARKYOALQGL-----GQNMSTIEEMAEKMTYAKDSS 1181
 1197 -----YLVKIHVKNGDHMKGARMLR 1218
 1182 DLPEESRRELLEQIADCCMRQSGSYHLATKKYQAGNKELKAMRALLK 1227

RESULT 14
 090G01 PRELIMINARY; PRT; 1653 AA.
 AC 090G01;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, last annotation update)
 DE HYPOTHETICAL 186.7 KDA PROTEIN (FRAGMENT).
 GN DKFZP434A163.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 RN NCBI_TaxId=9606;
 RX [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=TESTIS;
 RA Wambuit R., Heubner D., Mewes H.W., Gassenhuber J., Wiemann S.;
 RL Submitted (Aug-1999) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AL110218; CAB53678.1; -
 KW Hypothetical protein.
 FT NON-TER
 SQ SEQUENCE 1653 AA; 186734 MW; 670628A36CA762D4 CRC64;

Query Match 3.5%; Score 251.5; DB 4; Length 1653;
 Best Local Similarity 19.0%; Pred. No. 4.3e-08;
 Matches 243; Conservative 181; Mismatches 461; Indels 391; Gaps 58;

179 TKRITGCGMAENLXALG---GEDKMITVSNQEGDITROTQVSEPNXNQFELMKMDRT 235
 16 TSAVVCLOMPAYIIVFGIAEKGVRLANTKTKSSITGT-----ESVVSLLTNC 66
 236 SAAESNISVLLCKTLFLFLNLNPNPADLEFQDFGNIVC---YNNYGDGRVIGFSC 291
 67 SGKGLISGHADITVRYFP---DDEGSGESQGLVNHPCRPVALAMATNSIVAAG--C 119
 292 GHFVYISTHTGSGEIQARHKNLNLISVNSQTLNVATCGD--NCIKIDQVLDLKMXY 350
 120 DRKIYAYGKEGHMLQTFDYSRDPQREFFTAVSSGGQSVLGYSYDRLLRVFEMIRRSIW 179
 351 VILNIDENK-----GIATLSWTDDQGLALSTORSLHVFITKPLIADACSTRIA 402
 180 -----EEAKPEITINLITITALAMKRDGSRCLCVTLCGGVGF-----DCLLRSI 225

OY	597	TIKGIMWENPMOKGVIATDDDKVYTYVPHKDTIGCAVILLAGISIKYVFAH-----KPL	651
Db	615	-----KTGOIDRRETLSTNEDE--TNKSHLEFVEGLK-----NYPVNHFWDQSEPR	659
OY	652	LLY-----NGELTCOTOSGK-----VNNTYLS-TGFTLSNLDKXPD	687
Db	660	LFVCEAVOETPRSQPOGANG-----QRPDDGAGPAADVLLSPFISBEHFL--LHESRP-	712
OY	688	ELRPLMLAHNLM-----	702
Db	713	--RATSHSLIGMEVPYUFTYFRKPREADEVEPCGHHIPOMVSRRLPDLFGLEDCK	770
OY	703	-----SOMAMECRILNDEAMNLSAPACIHHNEVEPAIVYRRTGNY	744
Db	771	ATRDAMLHFSEFVTYIGMDERFKSTIKLISEAVENWARRCKVKTORDIVA-KVC--LGSM	827
OY	745	GIVVSLBOIKGIEDYNNLACHLAMPFTNDYMLAODL-YLASSCPALAEKRRDLOH-----	798
Db	828	GHANGARALRABERPELBARVAVIATQGLMEDAQLRK-----KRHLLAKFPQA	881
OY	799	---WDSALOQLAKHLAPDQIPFISKEVATOLEFAGDVNALAHYKGTGDKNKEHDACLA	855
Db	882	AGRWQOALQVAEHHDHRYHLRSTYHRYAGHLEASDCSRALSYEKESDT-----	929
OY	856	GVAQMSIRMGDIRGVNQALKHPSRYLAKRCCGAILENMKQFSEANQVLEKGIYDPAKSV	915
Db	930	-----HREVP-RMLSEDPUL-----ELIYVK-----	951
OY	916	YIRSKN---WAKYGDLLPHVSSPKIHLQYAKKEADGRYKEAVAYENAKOMOSVIRIY	971
Db	952	--MKDKTILMRWA-----QYL-----EEOGEMDAALHYELARPHFSLVATH	991
OY	972	LDHLNPEKAVNIYRETOSLDGAKMVARFLOLGDYGSALIOFLVNSKCNNEAFTLAQOH	1031
Db	992	C-FOGNVOKAAQIANTETGNAELASHYLARQYESQEEVGQAVHEFTYRAQAFKNAIRLCKEN-	1049
OY	1032	KMEIYADLIGSED-----TJNEDVQSIALTFEBEKKRYLQAGFPFL--CGQYSAL	1080
Db	1050	-----GDDQJLMLNALLSSPEDITEARY--EEKGVOMDRVMVLYHRAGHFSAL	1098
OY	1081	KHFLKCESSEEDNVAIEAMITFVQACADELTJNOLIDHLGENDGMPKDAKY--LPRYLMA	1138
Db	1099	ELAF---ATQGFVALQJLADIEDTSPRLALRCSOFI-----EHSQYBRAYELLIA	1148
OY	1139	LKQYREAAQTAIIITAREQSGAGNRYNAHDVLFESVAYELKSQKIKITPSEMAYNLMITHS--	1196
Db	1149	ARKYQOALQJLCL-----GOMNSTEEMAEKMTAVAKDSS	1181
OY	1197	-----YLYKTIHYANGSDHMGAMRLR	1218
Db	1182	DLPEBSRELLLEQIADCCMROGSHYHLATKRYQDAGKRLKAMRALIK	1227

Search completed: May 21, 2002, 14:53:17
Job time: 217 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 21, 2002, 14:50:35 ; Search time 38.73 Seconds
(without alignments)
3080.125 Million cell updates/sec

Title: US-09-729-653-2_COPY_1_1074
Perfect score: 1074
Sequence: 1 HSLIGRCSRGILGDNNAVAC.....LYFEGEKRYLQAKFFLLCG 1074

Scoring table:
Gapop 60.0 , Gapext 60.0

Searched: 747574 seqs, 111073796 residues

Word size : 0

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database : A_Geneseq_032802.*

```
1: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1980.DAT:*
2: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1981.DAT:*
3: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1982.DAT:*
4: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1983.DAT:*
5: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1984.DAT:*
6: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1985.DAT:*
7: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1986.DAT:*
8: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1987.DAT:*
9: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1988.DAT:*
10: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1989.DAT:*
11: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1990.DAT:*
12: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1991.DAT:*
13: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1992.DAT:*
14: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1993.DAT:*
15: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1994.DAT:*
16: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1995.DAT:*
17: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1996.DAT:*
18: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1997.DAT:*
19: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1998.DAT:*
20: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1999.DAT:*
21: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2000.DAT:*
22: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	367	34.2	502	22	AAW79303 Human protein SEQ
2	277	25.8	515	22	AAM80287 Human protein SEQ
3	80	7.4	138	22	AAU14751 Novel bone marrow
4	62	5.8	244	22	AAU14657 Novel bone marrow
5	32	3.0	72	22	AAU33013 Novel human secret
6	25	2.3	237	22	AAU33013 Novel human secret
7	19	1.8	109	22	AAU33123 Novel human secret
8	7	0.7	55	20	AAV41351 Human secreted pro
9	7	0.7	57	14	AAV41351 Human secreted pro
10	7	0.7	58	22	AAV41351 Human secreted pro
11	7	0.7	61	21	AAV41351 Human secreted pro

12	7	0.7	62	21	AAB39306 Human secreted pro
13	7	0.7	63	22	ABB44131 Peptide #11637 enc
14	7	0.7	63	22	ABB17718 Human nervous syst
15	7	0.7	63	22	ABB27013 Protein #9012 enco
16	7	0.7	63	22	AAW65154 Human bone marrow
17	7	0.7	63	22	AAW77859 Human bone marrow
18	7	0.7	63	22	AAW21759 Peptide #8193 enco
19	7	0.7	63	22	AAW38081 Peptide #12118 enc
20	7	0.7	65	21	AAW58410 Lung cancer associ
21	7	0.7	65	22	AAU4940 Protonibacterium
22	7	0.7	77	22	AAU57398 Protonibacterium
23	7	0.7	85	22	AAW76182 Human colon cancer
24	7	0.7	86	21	AAW26901 zea mays protein f
25	7	0.7	88	22	AAW89209 Human immune/haema
26	7	0.7	90	22	AAU33124 Novel human secret
27	7	0.7	98	22	AAU47185 Protonibacterium
28	7	0.7	98	22	AAW91472 Human immune/haema
29	7	0.7	106	22	AAV34835 Chlamydia pneumoni
30	7	0.7	110	21	AAW28487 zea mays protein f
31	7	0.7	114	20	AAV74083 Human prostate tum
32	7	0.7	114	20	AAV01147 Secreted protein e
33	7	0.7	117	21	AAW22616 zea mays protein f
34	7	0.7	119	21	AAV64673 Human 5' EST relat
35	7	0.7	130	21	AAW43607 Arabidopsis thalia
36	7	0.7	135	21	AAW43606 Arabidopsis thalia
37	7	0.7	137	21	AAW26900 zea mays protein f
38	7	0.7	140	21	AAW55757 Arabidopsis thalia
39	7	0.7	140	21	AAW61480 Arabidopsis thalia
40	7	0.7	140	22	AAW65756 Testis Enhanced Ge
41	7	0.7	141	22	AAW68476 Arabidopsis thalia
42	7	0.7	144	21	AAW15956 E. coli proliferat
43	7	0.7	144	22	AAU34747 E. coli cellular p
44	7	0.7	144	22	AAU35523 Haemophilus influe
45	7	0.7	144	22	AAU38506 Salmonella typhi c
46	7	0.7	144	22	AAW98241 Escherichia coli p
47	7	0.7	144	22	AAW98845 E. coli growth and
48	7	0.7	151	18	AAW27649 Secreted protein A
49	7	0.7	151	18	AAW44085 Human secreted pro
50	7	0.7	151	21	AAW10239 Human adult ovary
51	7	0.7	151	21	AAW13028 Arabidopsis thalia
52	7	0.7	151	21	AAW22615 zea mays protein f
53	7	0.7	151	21	AAW25116 Arabidopsis thalia
54	7	0.7	151	21	AAW42871 Arabidopsis thalia
55	7	0.7	153	22	AAW68216 Arabidopsis thalia
56	7	0.7	156	22	AAW52660 Protonibacterium
57	7	0.7	161	21	AAW13027 Arabidopsis thalia
58	7	0.7	161	21	AAW25115 Arabidopsis thalia
59	7	0.7	161	21	AAW42870 Arabidopsis thalia
60	7	0.7	163	20	AAV31928 Vernonia 1,3-beta-
61	7	0.7	170	21	AAW43605 Arabidopsis thalia
62	7	0.7	203	18	AAW28268 Amino acid sequenc
63	7	0.7	229	21	AAV44501 Glutathione-S-tran
64	7	0.7	232	21	AAW13026 Arabidopsis thalia
65	7	0.7	232	21	AAW42869 Arabidopsis thalia
66	7	0.7	235	22	AAU52773 Protonibacterium
67	7	0.7	249	13	AAW20144 Class B beta-lacta
68	7	0.7	249	13	AAW20573 Class B beta-lacta
69	7	0.7	249	13	AAW20574 Class B beta-lacta
70	7	0.7	258	13	AAW19279 Arabidopsis thalia
71	7	0.7	258	21	AAW48498 Arabidopsis thalia
72	7	0.7	269	21	AAW22220 Arabidopsis thalia
73	7	0.7	272	22	AAW66438 Human ATPase 30.
74	7	0.7	276	21	AAW48497 Arabidopsis thalia
75	7	0.7	277	22	AAW15556 Novel human diapo
76	7	0.7	284	21	AAW39054 Arabidopsis thalia
77	7	0.7	293	21	AAW22219 Arabidopsis thalia
78	7	0.7	295	21	AAW30200 Arabidopsis thalia
79	7	0.7	296	20	AAV36855 Protein Involved i
80	7	0.7	296	21	AAW35215 Enterococcus faeca
81	7	0.7	306	21	AAW22218 Arabidopsis thalia
82	7	0.7	310	21	AAW48496 Arabidopsis thalia
83	7	0.7	312	22	AAU27719 Human full-length
84	7	0.7	315	19	AAW98227 H. pylori GHP0 115

85	7	0.7	316	22	AAU48530	Proteinlactarium
86	7	0.7	322	22	ABBS5960	Drosophila melanog
87	7	0.7	328	21	AA616110	Arabidopsis thalia
88	7	0.7	330	21	AA639053	Arabidopsis thalia
89	7	0.7	334	21	AA614646	Arabidopsis thalia
90	7	0.7	336	22	AAU27891	Human contig polyp
91	7	0.7	338	15	AA65965	T. niveum GAPDH.
92	7	0.7	340	21	AAV75131	Neisseria meningit
93	7	0.7	341	22	ABG29988	Novel human diagno
94	7	0.7	344	21	AAV75130	Neisseria meningit
95	7	0.7	345	21	AA630199	Arabidopsis thalia
96	7	0.7	345	22	AAU33858	Staphylococcus aur
97	7	0.7	356	22	AAU00858	S. aureus D-alalan
98	7	0.7	359	17	AA688870	Sardinian tomato y
99	7	0.7	359	17	AA688871	Sardinian tomato y
100	7	0.7	359	17	AA688872	Sardinian tomato y
101	7	0.7	361	8	AA670562	Product of ORF 4 f
102	7	0.7	362	15	AA648694	G-protein coupled
103	7	0.7	362	17	AAW02666	G-protein coupled
104	7	0.7	364	21	AAV51230	Newcastle disease
105	7	0.7	373	20	AAV41692	Human PRO 363 prot
106	7	0.7	373	21	AA644248	Human PRO363 (UNQ3
107	7	0.7	373	21	AA633430	Human PRO363 prote
108	7	0.7	373	22	AAU12365	Human PRO363 polyp
109	7	0.7	373	22	AA648108	Human A236 polypep
110	7	0.7	373	22	AA648126	Mouse A236 polypep
111	7	0.7	373	22	AA648145	Mouse A236 variant
112	7	0.7	373	22	AA648146	Human A236 variant
113	7	0.7	373	22	AA648147	Human A236 variant
114	7	0.7	373	22	AA648148	Mouse A236 variant
115	7	0.7	373	22	AA648149	Mouse A236 variant
116	7	0.7	373	22	AA648150	Mouse A236 variant
117	7	0.7	373	22	AA65293	Human PRO363 prote
118	7	0.7	373	22	AA65293	Murine adipocytes-
119	7	0.7	379	22	AA652464	Mycobacterium ture
120	7	0.7	393	22	AAU36802	Staphylococcus aur
121	7	0.7	403	21	AA637514	Arabidopsis thalia
122	7	0.7	404	21	AA606781	Arabidopsis thalia
123	7	0.7	404	21	AA638369	Arabidopsis thalia
124	7	0.7	405	21	AA637513	Arabidopsis thalia
125	7	0.7	410	22	AA604102	Botulism toxin hea
126	7	0.7	411	22	AA639052	Arabidopsis thalia
127	7	0.7	426	22	AA673967	Human colon cancer
128	7	0.7	433	21	AA606780	Arabidopsis thalia
129	7	0.7	433	21	AA638368	Arabidopsis thalia
130	7	0.7	435	22	AA695505	Human protein sequ
131	7	0.7	439	21	AA606779	Arabidopsis thalia
132	7	0.7	439	21	AA638367	Arabidopsis thalia
133	7	0.7	443	22	AA614645	Arabidopsis thalia
134	7	0.7	443	22	AA620005	Arabidopsis 3-keto
135	7	0.7	446	21	AA614644	Arabidopsis thalia
136	7	0.7	446	22	ABG29986	Novel human diagno
137	7	0.7	447	21	AA637512	Arabidopsis thalia
138	7	0.7	448	20	AAV03773	Human T-box polype
139	7	0.7	450	21	AAV69997	Human receptor-ass
140	7	0.7	456	21	AA657017	Human prostate can
141	7	0.7	462	21	AA630198	Arabidopsis thalia
142	7	0.7	462	22	AA620004	Arabidopsis 3-keto
143	7	0.7	476	11	AA605599	BIV gag gene produ
144	7	0.7	480	22	AB64489	Drosophila melanog
145	7	0.7	490	22	AA620009	Brassica 3-ketoacy
146	7	0.7	510	20	AAV22201	Human extracellula
147	7	0.7	510	21	AAV54368	Human extracellula
148	7	0.7	510	21	AAV66646	Protein encoded by
149	7	0.7	510	22	AA603653	Membrane-bound pro
150	7	0.7	510	22	AA65169	Human extracellula
151	7	0.7	510	22	AA650955	Human PRO698 (UNQ3
152	7	0.7	525	22	AA604880	Human protease pro
153	7	0.7	525	22	AA694023	Human protein sequ
154	7	0.7	546	22	ABG23395	Novel human diagno
155	7	0.7	566	20	AAV23888	Auifex VF-5 DNA p
156	7	0.7	574	20	AAV00940	Agfex VF-5 DNA p
157	7	0.7	600	21	AAV44500	GST signal peptide

158	7	0.7	640	20	AAW87762	Maize lysine ketog
159	7	0.7	675	21	AA610163	ISS1/1 polymerase
160	7	0.7	684	21	AA610162	ISS1/1 polymerase
161	7	0.7	706	22	AA679588	Corynebacterium gl
162	7	0.7	745	22	AB658741	Drosophila melanog
163	7	0.7	750	21	AA631281	Arabidopsis thalia
164	7	0.7	752	17	AA697834	Kaposi's sarcoma a
165	7	0.7	752	17	AA697834	Kaposi's sarcoma a
166	7	0.7	753	22	AB667122	Drosophila melanog
167	7	0.7	760	21	AA631280	Arabidopsis thalia
168	7	0.7	769	21	AA631279	Arabidopsis thalia
169	7	0.7	777	22	AB660046	Drosophila melanog
170	7	0.7	798	22	AB662893	Drosophila melanog
171	7	0.7	799	21	AA640275	Human ORFX ORF39 p
172	7	0.7	811	22	AAU33166	Novel human secret
173	7	0.7	818	22	AB64426	Drosophila melanog
174	7	0.7	828	21	AAV44498	GST signal peptide
175	7	0.7	874	22	AA646696	Fowlpox virus DNA
176	7	0.7	889	22	AA668948	Arabidopsis thalia
177	7	0.7	906	22	AA681162	Mycobacterium tube
178	7	0.7	1019	22	AA692745	C glutamicum prote
179	7	0.7	1022	20	AA687761	Maize lysine ketog
180	7	0.7	1091	21	AA638615	Arabidopsis thalia
181	7	0.7	1096	21	AA638614	Arabidopsis thalia
182	7	0.7	1139	22	ABG22397	Novel human diagno
183	7	0.7	1140	21	AA638613	Arabidopsis thalia
184	7	0.7	1333	22	AB666754	Drosophila melanog
185	7	0.7	1346	21	AA639112	Arabidopsis thalia
186	7	0.7	1366	22	AB655570	Drosophila melanog
187	7	0.7	1390	21	AA639111	Arabidopsis thalia
188	7	0.7	1403	21	AA639110	Arabidopsis thalia
189	7	0.7	1430	22	AB658724	Drosophila melanog
190	7	0.7	1446	22	ABG29987	Novel human diagno
191	7	0.7	1455	22	AAW79120	Human protein SEQ
192	7	0.7	1612	19	AA65088	R. prowazekii S-la
193	7	0.7	1621	22	AAW25646	Human protein sequ
194	7	0.7	1668	22	AAU34158	Staphylococcus aur
195	7	0.7	1703	22	AB666223	Drosophila melanog
196	7	0.7	1953	22	AB662633	Drosophila melanog
197	7	0.7	2058	22	AA697070	Human polypeptide
198	7	0.7	2397	22	AAU36672	Staphylococcus aur
199	7	0.7	2435	22	AB660448	Drosophila melanog
200	6	0.6	7	21	AA619157	Peptide derived fr
201	6	0.6	9	21	AA601463	Minor histocompati
202	6	0.6	10	18	AAW40685	Peptide which bind
203	6	0.6	10	18	AAW40691	Peptide which bind
204	6	0.6	10	18	AAW40698	Peptide which bind
205	6	0.6	10	20	AAV43535	Leishmania mexican
206	6	0.6	10	21	AA601461	Minor histocompati
207	6	0.6	11	16	AA679901	Human FK-506 cytos
208	6	0.6	13	20	AAV23261	Murine apolipoprot
209	6	0.6	14	19	AAW40274	K. oxytoca R-speci
210	6	0.6	14	20	AAV23254	Murine apolipoprot
211	6	0.6	14	20	AAV23256	Apolipoprotein E d
212	6	0.6	14	20	AAV23257	Apolipoprotein E d
213	6	0.6	14	20	AAV23258	Apolipoprotein E d
214	6	0.6	14	20	AAV23259	Apolipoprotein E d
215	6	0.6	14	20	AAV23260	Steap/Gbeta bindin
216	6	0.6	15	20	AAV23260	Caenorhabditis ele
217	6	0.6	15	21	AA638099	Vaccine related MH
218	6	0.6	15	22	AA69455	Vaccine related MH
219	6	0.6	15	22	AA66557	Human cytomagalovl
220	6	0.6	15	22	AA66559	Human cytomagalovl
221	6	0.6	15	22	AA678616	Human zinc finger
222	6	0.6	16	22	AA699085	Vaccine related MH
223	6	0.6	17	16	AA65128	Cholera toxin B an
224	6	0.6	17	22	AA695553	Vaccine related MH
225	6	0.6	20	20	AAV12576	Human 5' EST secre
226	6	0.6	21	22	AA689179	HIV gp120 protein
227	6	0.6	21	22	AA689180	HIV gp120 protein
228	6	0.6	21	22	AA689181	HIV gp120 protein
229	6	0.6	22	22	AAW47407	Peptide #37 for iL
230	6	0.6	23	15	AA66107	pUR4421-encoded pe

231	AAV87538	6	0.6	24	21	Mature conotoxin p	304	6	0.6	53	22	AAU64777	Proponibacterium
232	AAV93093	6	0.6	25	10	CD4 anti-r-receptor	305	6	0.6	53	22	ABG09316	Novel human diagno
233	AAV09081	6	0.6	25	10	E6p mutated fragm	306	6	0.6	54	20	AAV29016	T. gondii immunoge
234	AAV20979	6	0.6	32	19	Human gliol fibril	307	6	0.6	54	22	AAU60158	Proponibacterium
235	AAV38503	6	0.6	34	21	Fragment of human	308	6	0.6	54	22	AAU25487	T. gondii immunoge
236	ABR38022	6	0.6	35	22	Peptide #5528 enco	309	6	0.6	55	22	AAE01648	Human secreted pro
237	ABR23256	6	0.6	35	22	Protein #5255 enco	310	6	0.6	56	21	AAU03395	Human secreted pro
238	AAV58649	6	0.6	35	22	Human brain expres	311	6	0.6	56	22	AAE01649	Human gene 18 enco
239	AAV71155	6	0.6	35	22	Human bone marrow	312	6	0.6	57	22	ABG24412	Novel human diagno
240	AAV31436	6	0.6	35	22	Peptide #5473 enco	313	6	0.6	57	22	AAU03197	Human polypeptide
241	AAV89533	6	0.6	36	22	Human immune/haema	314	6	0.6	58	21	AAV51413	Human secreted pro
242	ABG16492	6	0.6	38	22	Novel human diagno	315	6	0.6	58	22	AAU43788	Proponibacterium
243	ABG19318	6	0.6	38	22	Novel human diagno	316	6	0.6	58	22	ABG28760	Novel human diagno
244	AAV16889	6	0.6	40	21	Bacteriophage Dp-1	317	6	0.6	58	22	ABR30120	Peptide #2771 enco
245	ABR28224	6	0.6	40	22	Human peptide #875	318	6	0.6	58	22	ABR35292	Peptide #2798 enco
246	ABR33399	6	0.6	40	22	Peptide #905 enco	319	6	0.6	58	22	AAV96030	Human reproductive
247	ABR18858	6	0.6	40	22	Peptide #857 enco	320	6	0.6	58	22	AAU18878	Human bone marrow
248	AAV54184	6	0.6	40	22	Human brain expres	321	6	0.6	58	22	AAV56121	Novel prostate gla
249	AAV6578	6	0.6	40	22	Human bone marrow	322	6	0.6	58	22	AAV68493	Human bone marrow
250	AAV14451	6	0.6	40	22	Peptide #885 enco	323	6	0.6	58	22	AAV16301	Peptide #2735 enco
251	AAV26864	6	0.6	40	22	Peptide #801 enco	324	6	0.6	58	22	AAV04035	Peptide #2717 enco
252	AAV02178	6	0.6	40	22	Peptide #860 enco	325	6	0.6	59	22	AAU46459	Proponibacterium
253	AAV01464	6	0.6	41	21	Minor histocompati	326	6	0.6	59	22	ABG25720	Novel human diagno
254	AAV01465	6	0.6	41	21	Minor histocompati	327	6	0.6	60	22	AAU39131	Proponibacterium
255	AAV011979	6	0.6	41	21	Human polypeptide	328	6	0.6	60	22	AAV51250	Proponibacterium
256	ABR28150	6	0.6	42	22	Human peptide #801	329	6	0.6	60	22	ABG24749	Novel human diagno
257	ABR28157	6	0.6	42	22	Human peptide #818	330	6	0.6	60	22	ABR29102	Peptide #1753 enco
258	ABR30177	6	0.6	42	22	Peptide #2828 enco	331	6	0.6	60	22	ABR34265	Peptide #1771 enco
259	ABR33325	6	0.6	42	22	Peptide #831 enco	332	6	0.6	60	22	ABR19698	Protein #1697 enco
260	ABR33342	6	0.6	42	22	Peptide #848 enco	333	6	0.6	60	22	AAU21193	Human novel foetal
261	ABR35343	6	0.6	42	22	Peptide #2849 enco	334	6	0.6	60	22	AAV55056	Human brain expres
262	ABR18785	6	0.6	42	22	Protein #784 enco	335	6	0.6	60	22	AAV67448	Human bone marrow
263	ABR18802	6	0.6	42	22	Protein #801 enco	336	6	0.6	60	22	AAV15274	Peptide #1708 enco
264	ABR20784	6	0.6	42	22	Protein #2783 enco	337	6	0.6	60	22	AAV27740	Peptide #1771 enco
265	AAV54128	6	0.6	42	22	Human brain expres	338	6	0.6	60	22	AAV03019	Peptide #1701 enco
266	AAV56175	6	0.6	42	22	Human brain expres	339	6	0.6	61	21	AAV03138	Human secreted pro
267	AAV65504	6	0.6	42	22	Human bone marrow	340	6	0.6	61	22	AAV52545	Proponibacterium
268	AAV65521	6	0.6	42	22	Human bone marrow	341	6	0.6	61	22	ABR40905	Peptide #841 enco
269	AAV68547	6	0.6	42	22	Human bone marrow	342	6	0.6	61	22	ABR25034	Protein #7033 enco
270	AAV14377	6	0.6	42	22	Peptide #811 enco	343	6	0.6	61	22	AAV61764	Human brain expres
271	AAV14394	6	0.6	42	22	Peptide #828 enco	344	6	0.6	61	22	AAV74562	Human bone marrow
272	AAV16354	6	0.6	42	22	Peptide #2788 enco	345	6	0.6	61	22	AAV20342	Peptide #6776 enco
273	AAV26790	6	0.6	42	22	Peptide #827 enco	346	6	0.6	61	22	AAV34674	Peptide #8711 enco
274	AAV26807	6	0.6	42	22	Peptide #844 enco	347	6	0.6	61	22	AAV76757	Human colon cancer
275	AAV28851	6	0.6	42	22	Peptide #2888 enco	348	6	0.6	62	22	AAV43437	Proponibacterium
276	AAV02121	6	0.6	42	22	Peptide #803 enco	349	6	0.6	62	22	AAV46812	Proponibacterium
277	AAV04092	6	0.6	42	22	Peptide #2774 enco	350	6	0.6	62	22	ABG09165	Novel human diagno
278	ABR40314	6	0.6	43	22	Peptide #7820 enco	351	6	0.6	63	22	ABR32216	Peptide #4867 enco
279	ABR24715	6	0.6	43	22	Protein #6714 enco	352	6	0.6	63	22	ABR40694	Peptide #8200 enco
280	AAV61114	6	0.6	43	22	Human brain expres	353	6	0.6	63	22	ABR17324	Human nervous syst
281	AAV73822	6	0.6	43	22	Human bone marrow	354	6	0.6	63	22	AAV61556	Human brain expres
282	AAV20114	6	0.6	43	22	Peptide #5548 enco	355	6	0.6	63	22	AAV74344	Human bone marrow
283	AAV34008	6	0.6	43	22	Peptide #8045 enco	356	6	0.6	63	22	AAV18426	Peptide #4860 enco
284	AAV58848	6	0.6	44	21	Arbidopsis thalia	357	6	0.6	63	22	AAV34457	Peptide #8494 enco
285	ABG24649	6	0.6	44	22	Novel human diagno	358	6	0.6	63	22	AAV06024	Peptide #4706 enco
286	AAV85297	6	0.6	45	22	Human immune/haema	359	6	0.6	64	22	AAV95255	Human reproductive
287	AAV38172	6	0.6	46	21	Gene 3 human secre	360	6	0.6	65	21	AAV59870	Arbidopsis thalia
288	AAV38173	6	0.6	46	21	Human secreted pro	361	6	0.6	65	22	AAV47323	Proponibacterium
289	AAV32227	6	0.6	48	22	Novel human secre	362	6	0.6	66	21	AAV76173	Human secreted pro
290	ABR23035	6	0.6	49	22	Protein #5034 enco	363	6	0.6	66	22	AAV61207	Proponibacterium
291	AAE04847	6	0.6	49	22	Human SGP014 phosp	364	6	0.6	66	22	AAV66353	Proponibacterium
292	AAV74753	6	0.6	50	19	Human secreted pro	365	6	0.6	66	22	AAV007539	Human polypeptide
293	AAU42829	6	0.6	50	22	Proponibacterium	366	6	0.6	67	15	AAV46071	T-complex like pro
294	AAV55376	6	0.6	51	22	Proponibacterium	367	6	0.6	67	19	AAV20950	Human presentilin I
295	AAV52630	6	0.6	52	22	Proponibacterium	368	6	0.6	67	20	AAV12959	Human 5' EST secre
296	AAV67169	6	0.6	52	22	Proponibacterium	369	6	0.6	67	22	ABR27655	Human peptide #306
297	ABG06727	6	0.6	52	22	Novel human diagno	370	6	0.6	67	22	ABR28235	Human peptide #886
298	ABG08764	6	0.6	52	22	Novel human diagno	371	6	0.6	67	22	ABR32825	Peptide #331 enco
299	ABR15901	6	0.6	52	22	Human nervous syst	372	6	0.6	67	22	ABR33410	Peptide #916 enco
300	AAU20569	6	0.6	52	22	Human secreted pro	373	6	0.6	67	22	ABR37859	Peptide #5365 enco
301	AAV02496	6	0.6	53	21	Human secreted pro	374	6	0.6	67	22	ABR18307	Protein #306 enco
302	AAV5232	6	0.6	53	21	Human 5' EST relat	375	6	0.6	67	22	ABR18869	Protein #868 enco
303	AAU51376	6	0.6	53	22	Proponibacterium	376	6	0.6	67	22	ABR22960	Protein #4959 enco

377	6	0.6	67	22	AAW53631	Human brain expres	450	6	0.6	82	20	AAV35247	Chlamydia pneumonia
378	6	0.6	67	22	AAW54195	Human brain expres	451	6	0.6	82	21	AAW52034	Human secreted pro
379	6	0.6	67	22	AAW58490	Human brain expres	452	6	0.6	82	21	AAW07439	Arabidopsis thalia
380	6	0.6	67	22	AAW66011	Human bone marrow	453	6	0.6	82	21	AAW34762	Arabidopsis thalia
381	6	0.6	67	22	AAW66589	Human bone marrow	454	6	0.6	83	21	AAW32727	Zea mays protein f
382	6	0.6	67	22	AAW70990	Human bone marrow	455	6	0.6	84	21	AAW25303	Zea mays protein f
383	6	0.6	67	22	AAW89672	Human Immune/haema	456	6	0.6	85	18	AAW55416	H. pylori ORF hpi1
384	6	0.6	67	22	AAW13880	Peptide #314 encod	457	6	0.6	85	22	AAW53552	Proionibacterium
385	6	0.6	67	22	AAW14462	Peptide #896 encod	458	6	0.6	85	22	AAW09985	Human polypeptide
386	6	0.6	67	22	AAW18761	Peptide #5195 encod	459	6	0.6	86	22	AAW46479	B. brevis tyrocid
387	6	0.6	67	22	AAW26287	Peptide #324 encod	460	6	0.6	87	22	AAW52459	Proionibacterium
388	6	0.6	67	22	AAW26875	Peptide #912 encod	461	6	0.6	87	22	AAW52497	Proionibacterium
389	6	0.6	67	22	AAW31266	Peptide #5303 encod	462	6	0.6	88	21	AAW43149	Human ORFX ORF2913
390	6	0.6	67	22	AAW01623	Peptide #305 encod	463	6	0.6	88	22	AAW03843	Human musculoskele
391	6	0.6	67	22	AAW02189	Peptide #871 encod	464	6	0.6	88	22	AAW94895	Human reproductiv
392	6	0.6	68	22	AAW83810	Human Immune/haema	465	6	0.6	89	21	AAW26659	Arabidopsis thalia
393	6	0.6	68	22	AAW09990	Human polypeptide	466	6	0.6	89	22	AAW70551	Drosophila melanog
394	6	0.6	69	11	AAW05873	N-terminal sequenc	467	6	0.6	89	22	AAW87108	Human immune/haema
395	6	0.6	69	21	AAW34050	Human secreted pro	468	6	0.6	91	18	AAW27796	Iron (III) dicitra
396	6	0.6	69	21	AAW19806	Arabidopsis thalia	469	6	0.6	91	21	AAW01837	Human secreted pro
397	6	0.6	69	21	AAW33703	Arabidopsis thalia	470	6	0.6	92	20	AAW36565	Fragment of human
398	6	0.6	69	22	AAW32313	Peptide #4964 encod	471	6	0.6	92	21	AAW56482	Human prostate can
399	6	0.6	69	22	AAW22824	Human prostate can	472	6	0.6	93	21	AAW91561	Human secreted pro
400	6	0.6	69	22	AAW96130	Human reproductiv	473	6	0.6	93	21	AAW91679	Human secreted pro
401	6	0.6	69	22	AAW58230	Human brain expres	474	6	0.6	94	15	AAW2442	Guanylate cyclase
402	6	0.6	69	22	AAW07169	Human polypeptide	475	6	0.6	94	16	AAW69810	Expression product
403	6	0.6	69	22	AAW30994	Peptide #5031 encod	476	6	0.6	94	22	AAW08920	Human NOVX1 protei
404	6	0.6	70	21	AAW42091	Human ORFX ORF1855	477	6	0.6	94	22	AAW91167	Human immune/haema
405	6	0.6	70	22	AAW60025	Proionibacterium	478	6	0.6	95	11	AAW04198	F gene of simian 1
406	6	0.6	70	22	AAW29249	Peptide #1900 encod	479	6	0.6	95	18	AAW20877	H. pylori cytoplas
407	6	0.6	70	22	AAW34412	Peptide #1918 encod	480	6	0.6	95	22	AAW20451	Arabidopsis thalia
408	6	0.6	70	22	AAW19824	Protein #1823 encod	481	6	0.6	95	22	AAW39861	Proionibacterium
409	6	0.6	70	22	AAW55203	Human brain expres	482	6	0.6	96	21	AAW56643	Partial peptide fr
410	6	0.6	70	22	AAW67597	Human bone marrow	483	6	0.6	97	22	AAW313754	Novel human diagno
411	6	0.6	70	22	AAW15407	Peptide #1841 encod	484	6	0.6	99	12	AAW92035	Sequence encoded 1
412	6	0.6	70	22	AAW27895	Peptide #1932 encod	485	6	0.6	99	22	AAW44904	Proionibacterium
413	6	0.6	70	22	AAW03168	Peptide #1850 encod	486	6	0.6	99	22	AAW47407	Proionibacterium
414	6	0.6	71	22	AAW74469	Human colon cancer	487	6	0.6	99	22	AAW17307	Novel human diagno
415	6	0.6	72	21	AAW87539	Conotoxin peptide	488	6	0.6	100	19	AAW64533	Novel human diagno
416	6	0.6	73	22	AAW63961	Proionibacterium	489	6	0.6	100	22	AAW65880	Human C4 fragment
417	6	0.6	73	22	AAW61092	Novel human diagno	490	6	0.6	100	22	AAW50597	Drosophila melanog
418	6	0.6	73	22	AAW02183	Human polypeptide	491	6	0.6	100	22	AAW29572	Proionibacterium
419	6	0.6	74	21	AAW33702	Arabidopsis thalia	492	6	0.6	101	22	AAW34862	Novel human secret
420	6	0.6	74	22	AAW87509	Human Immune/haema	493	6	0.6	101	21	AAW21731	Human secreted pro
421	6	0.6	75	21	AAW19653	Arabidopsis thalia	494	6	0.6	101	22	AAW56031	Arabidopsis thalia
422	6	0.6	75	21	AAW00665	Human secreted pro	495	6	0.6	101	22	AAW32073	Proionibacterium
423	6	0.6	75	22	AAW52971	Proionibacterium	496	6	0.6	101	22	AAW04124	Novel human secret
424	6	0.6	76	20	AAW73859	Human prostate tum	497	6	0.6	101	22	AAW01825	Human gene 15 encod
425	6	0.6	76	21	AAW14414	Saccharomyces cere	498	6	0.6	101	22	AAW49132	Staphylococcus aur
426	6	0.6	76	22	AAW21167	Human novel foetal	499	6	0.6	102	21	AAW56617	Human prostate can
427	6	0.6	76	22	AAW17906	Novel human respir	500	6	0.6	102	21	AAW40808	Human ORFX ORF572
428	6	0.6	76	22	AAW75570	Human colon cancer	501	6	0.6	102	21	AAW42786	Human ORFX ORF2550
429	6	0.6	77	22	AAW87354	Human Immune/haema	502	6	0.6	102	21	AAW44007	Zea mays protein f
430	6	0.6	78	18	AAW55272	H. pylori ORF 05ce	503	6	0.6	103	6	AAW03440	Sequence of sub-un
431	6	0.6	78	22	AAW29525	Peptide #2176 encod	504	6	0.6	103	17	AAW04857	Synthetic cholera
432	6	0.6	78	22	AAW34709	Peptide #2215 encod	505	6	0.6	103	17	AAW06606	Cholera toxin B su
433	6	0.6	78	22	AAW20121	Protein #2120 encod	506	6	0.6	103	17	AAW06607	Cholera toxin B su
434	6	0.6	78	22	AAW55505	Human brain expres	507	6	0.6	103	19	AAW080808	Amino acid sequenc
435	6	0.6	78	22	AAW67890	Human bone marrow	508	6	0.6	103	21	AAW06063	Caenorhabditis ele
436	6	0.6	78	22	AAW15708	Peptide #2142 encod	509	6	0.6	103	21	AAW06130	Caenorhabditis ele
437	6	0.6	78	22	AAW28212	Peptide #2249 encod	510	6	0.6	103	21	AAW18769	Zea mays protein f
438	6	0.6	78	22	AAW03443	Peptide #2125 encod	511	6	0.6	103	22	AAW67871	Proionibacterium
439	6	0.6	79	22	AAW064318	Proionibacterium	512	6	0.6	103	22	AAW62362	V. cholera cholera
440	6	0.6	79	22	AAW00177	Human polypeptide	513	6	0.6	103	22	AAW62364	V. cholera cholera
441	6	0.6	80	21	AAW65667	C. elegans insulin	514	6	0.6	103	22	AAW62365	V. cholera cholera
442	6	0.6	80	22	AAW45074	Proionibacterium	515	6	0.6	103	22	AAW62367	V. cholera cholera
443	6	0.6	81	21	AAW00313	Arabidopsis thalia	516	6	0.6	103	22	AAW62370	V. cholera cholera
444	6	0.6	81	21	AAW00313	Human secreted pro	517	6	0.6	104	17	AAW04239	Human FRBP12. Hom
445	6	0.6	81	22	AAW49771	Proionibacterium	518	6	0.6	104	19	AAW56026	FK506-binding prot
446	6	0.6	81	22	AAW61059	Human polypeptide	519	6	0.6	104	21	AAW25021	Arabidopsis thalia
447	6	0.6	81	22	AAW01037	Human polypeptide	520	6	0.6	104	21	AAW40961	Zea mays protein f
448	6	0.6	81	22	AAW64202	Subtilopeptidase 9	521	6	0.6	104	21	AAW00653	Human secreted pro
449	6	0.6	81	22	AAW02043	Human secreted pro	522	6	0.6	104	22	AAW15388	Novel human diagno

523	6	0.6	104	22	ABG16383	Novel human diagno	596	6	0.6	120	22	AAU21843	Novel human neopla
524	6	0.6	105	22	AAU53993	Propionibacterium	597	6	0.6	121	22	AAU57757	Propionibacterium
525	6	0.6	105	22	AAU64367	Propionibacterium	598	6	0.6	121	22	AAE10605	Human macrophage-e
526	6	0.6	105	22	ABG10410	Novel human diagno	599	6	0.6	121	22	AAU08295	Human PML-like pro
527	6	0.6	105	22	ABG15359	Novel human diagno	600	6	0.6	121	22	AAU31861	Novel human secret
528	6	0.6	105	22	ABG15350	Novel human diagno	601	6	0.6	121	22	AAU32910	C glutamicum prote
529	6	0.6	105	22	AAU16284	Novel human secret	602	6	0.6	121	22	AAU94940	Human protein sequ
530	6	0.6	105	22	AAU16284	V. cholera cholera	603	6	0.6	122	22	AAU36408	Pseudomonas aerugi
531	6	0.6	106	12	AAU11725	Amino acids 1-106	604	6	0.6	122	22	AAU80854	Human haematologic
532	6	0.6	106	12	AAU79307	Human protein SEQ	605	6	0.6	122	22	AAU81313	Human haematologic
533	6	0.6	107	12	AAU11581	Macrocyclic FK-506	606	6	0.6	122	22	AAU81872	Human haematologic
534	6	0.6	107	16	AAU79857	Bovine FK-506 immu	607	6	0.6	122	22	AAU00622	Human polypeptide
535	6	0.6	107	16	AAU79857	MDV L1 protein. M	608	6	0.6	123	18	AAU28331	Staphylococcus aur
536	6	0.6	107	18	AAU21691	Human FK-binding p	609	6	0.6	123	22	ABU16477	Human nervous syst
537	6	0.6	107	18	AAU21691	FK506 binding prot	610	6	0.6	123	22	AAU31170	Novel human secret
538	6	0.6	107	19	AAU37470	Marek's disease vi	611	6	0.6	123	22	AAU86760	Human immune/haema
539	6	0.6	107	22	AAU28592	Amino acids 1-107	612	6	0.6	124	13	AAU28831	B subunit of CT.
540	6	0.6	107	21	AAU93339	An Escherichia col	613	6	0.6	124	17	AAU06805	Cholera toxin B su
541	6	0.6	107	22	AAU61584	Propionibacterium	614	6	0.6	124	21	AAU96872	Plant-optimized V.
542	6	0.6	107	22	ABU17608	Human nervous syst	615	6	0.6	124	22	AAU65992	Cholera toxin B su
543	6	0.6	108	13	AAU20561	FKBP. Homo sapien	616	6	0.6	124	22	AAU62359	V. cholera strain
544	6	0.6	108	13	AAU20561	FK-506 binding pro	617	6	0.6	124	22	AAU62361	V. cholera cholera
545	6	0.6	108	14	AAU43892	FK506 binding prot	618	6	0.6	124	22	AAU62363	V. cholera cholera
546	6	0.6	108	19	AAU61220	Streptococcus pneu	619	6	0.6	124	22	AAU62366	V. cholera cholera
547	6	0.6	108	21	AAU41491	Human ORFX ORF1255	620	6	0.6	124	22	AAU62368	V. cholera cholera
548	6	0.6	108	21	AAU44401	Human FKBP-12 cyto	621	6	0.6	125	12	AAU33017	Novel human secret
549	6	0.6	108	22	AAU04561	Human G-protein co	622	6	0.6	126	12	AAU12630	GtB.1/CTB chimeri
550	6	0.6	108	22	AAU31551	Novel human secret	623	6	0.6	126	21	AAU01793	Human secreted pro
551	6	0.6	109	22	AAU94914	Human reproductive	624	6	0.6	126	22	ABG14324	Novel human diagno
552	6	0.6	109	22	AAU603895	Human secreted pro	625	6	0.6	126	22	AAU00132	Human polypeptide
553	6	0.6	109	22	ABG06434	Novel human diagno	626	6	0.6	126	22	AAU73350	Human colon cancer
554	6	0.6	109	22	ABG24079	Novel human diagno	627	6	0.6	127	20	AAU73985	Human prostate tum
555	6	0.6	110	22	AAU93389	Human polypeptide.	628	6	0.6	127	21	AAU26604	Arabidopsis thalia
556	6	0.6	110	22	AAU01581	Human polypeptide	629	6	0.6	127	22	ABG03345	Novel human diagno
557	6	0.6	111	21	AAU53250	Human colon cancer	630	6	0.6	127	22	ABG13520	Novel human diagno
558	6	0.6	112	22	AAU41630	Propionibacterium	631	6	0.6	128	18	AAU08427	Vector expressing
559	6	0.6	112	22	AAU003128	Human polypeptide	632	6	0.6	128	20	AAU30055	Amino acid sequenc
560	6	0.6	113	22	AAU72085	Glycine max RAD51	633	6	0.6	128	22	AAU86553	Human immune/haema
561	6	0.6	114	21	AAU26658	Arabidopsis thalia	634	6	0.6	128	22	AAU01856	Human polypeptide
562	6	0.6	114	22	AAU59848	Propionibacterium	635	6	0.6	129	21	AAU45369	Gene 35 human secr
563	6	0.6	114	22	AAU005102	Human polypeptide	636	6	0.6	129	21	AAU45370	Human secreted pro
564	6	0.6	115	20	AAU42459	Human guanylin pre	637	6	0.6	129	22	AAU005272	Human polypeptide
565	6	0.6	115	20	AAU42460	Human guanylin pre	638	6	0.6	130	20	AAU29523	Human lung tumour
566	6	0.6	115	21	AAU12953	Arabidopsis thalia	639	6	0.6	130	20	AAU21597	Human secreted pro
567	6	0.6	115	21	AAU56949	Arabidopsis thalia	640	6	0.6	130	21	AAU58516	Lung cancer associ
568	6	0.6	115	21	AAU57269	Arabidopsis thalia	641	6	0.6	130	21	AAU44408	Human lung tumour
569	6	0.6	115	21	AAU02699	Human secreted pro	642	6	0.6	130	22	AAU13749	Human lung tumour-
570	6	0.6	115	22	AAU78557	Human protein SEQ	643	6	0.6	131	11	AAU04825	LTR-CTB fusion pro
571	6	0.6	115	22	AAU91675	Human immune/haema	644	6	0.6	131	21	AAU51234	Arabidopsis thalia
572	6	0.6	115	22	AAU92207	C glutamicum prote	645	6	0.6	131	21	AAU56948	Arabidopsis thalia
573	6	0.6	116	19	AAU52987	Homo sapiens clone	646	6	0.6	131	21	AAU57268	Arabidopsis thalia
574	6	0.6	116	20	AAU21592	Human secreted pro	647	6	0.6	131	21	AAU03893	Human secreted pro
575	6	0.6	116	21	AAU53345	Human colon cancer	648	6	0.6	131	22	AAU70623	Drosophila melanog
576	6	0.6	116	21	AAU26942	Zea mays protein f	649	6	0.6	131	22	AAU25454	Novel human diagno
577	6	0.6	117	10	AAU82088	Human haematologic	650	6	0.6	131	22	AAU66714	C glutamicum phosp
578	6	0.6	117	10	AAU90153	Sequence of hepati	651	6	0.6	132	21	AAU32654	Zea mays protein f
579	6	0.6	117	21	AAU92036	Sequence encoded i	652	6	0.6	132	22	ABU19027	Novel human diagno
580	6	0.6	117	21	AAU18080	Pinus radiata squa	653	6	0.6	132	22	AAU01519	Human polypeptide
581	6	0.6	118	11	AAU04163	Cholera toxin B-su	654	6	0.6	133	21	AAU10853	Arabidopsis thalia
582	6	0.6	118	20	AAU86250	Matize PTE partial	655	6	0.6	133	22	AAU84935	Human immune/haema
583	6	0.6	118	22	AAU63059	Propionibacterium	656	6	0.6	133	13	AAU20155	Segment of human C
584	6	0.6	118	22	ABU10853	Human ovarian and/	657	6	0.6	134	21	AAU04653	Arabidopsis thalia
585	6	0.6	118	22	AAU94530	Human reproductive	658	6	0.6	134	21	AAU21826	Arabidopsis thalia
586	6	0.6	118	22	AAU00762	Human polypeptide	659	6	0.6	134	21	AAU57552	Arabidopsis thalia
587	6	0.6	119	21	AAU34376	Arabidopsis thalia	660	6	0.6	134	21	AAU57659	Arabidopsis thalia
588	6	0.6	119	22	AAU01960	Human polypeptide	661	6	0.6	134	22	AAU81659	S. epidermidis ope
589	6	0.6	119	22	AAU02547	Human polypeptide	662	6	0.6	135	10	AAU90289	Sequence of a fusi
590	6	0.6	119	22	AAU06550	Human polypeptide	663	6	0.6	135	21	AAU14523	Arabidopsis thalia
591	6	0.6	120	19	AAU74835	Human secreted pro	664	6	0.6	135	21	AAU20966	Arabidopsis thalia
592	6	0.6	120	22	AAU26657	Arabidopsis thalia	665	6	0.6	135	21	AAU59173	CDA fragment bindi
593	6	0.6	120	22	AAU23179	Novel human enzyme	666	6	0.6	135	22	AAU02870	Human polypeptide
594	6	0.6	120	22	AAU01492	Human polypeptide	667	6	0.6	136	21	AAU58752	Breast and ovarian
595	6	0.6	120	22	AAU02771	Human NKp30 recept	668	6	0.6	136	21	AAU10414	Arabidopsis thalia

659	6	0.6	136	21	AA623102	Arabidopsis thalia	742	6	0.6	150	22	AAU23287	Novel human enzyme
670	6	0.6	136	21	AA63885	Arabidopsis thalia	743	6	0.6	150	22	AA672988	Olfactory receptor
671	6	0.6	136	22	ABBI0412	Human cDNA SEQ ID	744	6	0.6	151	12	AA613327	HEL epididymis-spe
672	6	0.6	136	22	AA689699	Human immune/haema	745	6	0.6	151	18	AA627292	Human H1075-1 secr
673	6	0.6	136	22	AA691793	Human immune/haema	746	6	0.6	151	19	AA681777	Human HEL protein
674	6	0.6	136	22	AA605854	Human polypeptide	747	6	0.6	151	19	AA603731	Human secreted pro
675	6	0.6	136	22	AAU16281	Human novel secret	748	6	0.6	151	21	AA693597	Protein encoded by
676	6	0.6	137	22	AB660526	Drosophila melanog	749	6	0.6	151	22	AA616399	Human nervous syst
677	6	0.6	137	22	AB629219	Novel human diago	750	6	0.6	151	22	AA678556	Human protein
678	6	0.6	137	22	AA694549	Human reproductive	751	6	0.6	152	19	AA648271	Rat ninturin 1. R
679	6	0.6	137	22	AA603957	Human polypeptide	752	6	0.6	152	19	AA648272	Human ninturin 1.
680	6	0.6	138	18	AA635844	Human CD4 for use	753	6	0.6	152	22	AA643525	Protonibacterium
681	6	0.6	138	18	AA608429	pCANTAB-AP-FKBP.	754	6	0.6	152	22	AA699859	Physcomitrella pat
682	6	0.6	138	21	AA644858	Zea mays protein f	755	6	0.6	152	22	AA692981	C glutamicum prote
683	6	0.6	138	21	AA656707	Amino acid sequenc	756	6	0.6	155	16	AA679152	E.coli ribonucleas
684	6	0.6	138	22	AAU57401	Protonibacterium	757	6	0.6	155	22	AA611967	Human secreted pro
685	6	0.6	139	20	AA635589	C. pneumoniae prot	758	6	0.6	155	22	AAU30231	Novel human secret
686	6	0.6	139	21	AA614522	Arabidopsis thalia	759	6	0.6	155	22	AA679540	Human protein SEQ
687	6	0.6	139	21	AA620965	Arabidopsis thalia	760	6	0.6	155	22	AA679541	Human protein SEQ
688	6	0.6	139	21	AA651233	Arabidopsis thalia	761	6	0.6	155	22	AA600453	Human polypeptide
689	6	0.6	139	22	AB609412	Novel human diago	762	6	0.6	156	22	AAU36521	Pseudomonas aerugi
690	6	0.6	139	22	AB641127	Peptide #8633 enco	763	6	0.6	156	22	AAU39788	Protonibacterium
691	6	0.6	139	22	AA661983	Human brain expres	764	6	0.6	157	21	AA644006	Zea mays protein f
692	6	0.6	139	22	AA674787	Human bone marrow	765	6	0.6	158	21	AA618768	Zea mays protein f
693	6	0.6	139	22	AA612894	Human polypeptide	766	6	0.6	158	21	AA619100	Novel human diago
694	6	0.6	139	22	AA634903	Peptide #8940 enco	767	6	0.6	159	18	AA644125	Streptococcus pneu
695	6	0.6	140	21	AA614521	Arabidopsis thalia	768	6	0.6	160	18	AA611225	S. pneumoniae taga
696	6	0.6	140	21	AA620964	Arabidopsis thalia	769	6	0.6	160	18	AA614722	C. hirsutus GDP.
697	6	0.6	140	22	AAU42645	Protonibacterium	770	6	0.6	160	20	AA673637	S. pneumoniae lacC
698	6	0.6	141	17	AA694409	Met-KGF delta-N23/	771	6	0.6	160	21	AA626941	Streptococcus pneu
699	6	0.6	141	17	AA693491	Keratinocyte growt	772	6	0.6	160	22	AA666423	Enterococcus faeca
700	6	0.6	141	19	AA698794	H. pylori GHP 126	773	6	0.6	161	20	AA600081	Enterococcus faeca
701	6	0.6	141	19	AA661433	Keratinocyte growt	774	6	0.6	161	20	AA690059	Pseudomonas sp. WF
702	6	0.6	141	20	AA628500	hbp protein. Legi	775	6	0.6	161	21	AA697828	Protonibacterium
703	6	0.6	141	21	AA616393	Eucahyptus thalia	776	6	0.6	161	22	AA666010	Human protein sequ
704	6	0.6	141	21	AA651232	Arabidopsis thalia	777	6	0.6	161	22	AA625783	Human H1075-1 secr
705	6	0.6	141	22	AAU35696	Helicobacter pylor	778	6	0.6	162	18	AA627291	Protein encoded by
706	6	0.6	141	22	AAU35878	Helicobacter pylor	779	6	0.6	162	18	AA636949	Protonibacterium
707	6	0.6	142	18	AA628124	Staphylococcus aur	780	6	0.6	162	22	AAU59012	Protonibacterium
708	6	0.6	142	18	AA628084	Threonine dehydrat	781	6	0.6	162	22	AAU63417	Human quanylate ki
709	6	0.6	142	22	AAU47085	Protonibacterium	782	6	0.6	163	22	AB612270	Human quanylate ki
710	6	0.6	143	22	AB600477	Novel human diago	783	6	0.6	163	22	AA625236	Eucahyptus grandis
711	6	0.6	143	22	AB628995	Novel human diago	784	6	0.6	164	21	AA616379	Human PRO230 amino
712	6	0.6	144	21	AA635445	Human colon cancer	785	6	0.6	164	21	AA688572	PRO230 tubulinlter
713	6	0.6	144	21	AA600123	HELIOP53 tumour su	786	6	0.6	164	21	AA688376	Human PRO230 polyI
714	6	0.6	144	21	AA609966	Arabidopsis thalia	787	6	0.6	164	22	AAU12328	Human PRO230. Hom
715	6	0.6	144	21	AA650454	Arabidopsis thalia	788	6	0.6	164	22	AA620339	Human PRO230. Hom
716	6	0.6	144	21	AA654790	Arabidopsis thalia	789	6	0.6	164	22	AA668597	PRO230. Homo sapi
717	6	0.6	144	22	AAU54933	Protonibacterium	790	6	0.6	164	22	AA680214	Human PRO230 prote
718	6	0.6	144	22	AA697910	Human GTP-binding	791	6	0.6	165	19	AA611071	H. pylori ORF 09ge
719	6	0.6	145	18	AA627812	Staphylococcus aur	792	6	0.6	165	21	AA657227	Human prostate can
720	6	0.6	145	20	AA637604	Protein which is s	793	6	0.6	165	21	AA6312085	Human secreted pro
721	6	0.6	145	21	AA632519	S. lavendulae Mmcf	794	6	0.6	165	21	AA609965	Arabidopsis thalia
722	6	0.6	146	21	AB610168	Human cDNA SEQ ID	795	6	0.6	165	21	AA650453	Arabidopsis thalia
723	6	0.6	146	21	AA656771	Human prostate can	796	6	0.6	165	22	AAU48220	Protonibacterium
724	6	0.6	146	22	AA675237	Human colon cancer	797	6	0.6	166	21	AA603986	Human secreted pro
725	6	0.6	147	22	AAU14415	Human novel protei	798	6	0.6	166	22	AAU56780	Protonibacterium
726	6	0.6	148	13	AA620153	N-terminal portion	799	6	0.6	166	22	AAU58218	Protonibacterium
727	6	0.6	148	22	AAU29520	Novel human secret	800	6	0.6	166	22	AAU62962	Protonibacterium
728	6	0.6	149	21	AA657795	Arabidopsis thalia	801	6	0.6	166	22	AB607311	Novel human diago
729	6	0.6	149	22	AB618726	Novel human diago	802	6	0.6	167	21	AA616491	Arabidopsis thalia
730	6	0.6	149	22	AB625600	Novel human diago	803	6	0.6	167	21	AA637995	Arabidopsis thalia
731	6	0.6	149	22	AA690245	Human immune/haema	804	6	0.6	167	21	AA638326	Arabidopsis thalia
732	6	0.6	149	22	AA603624	Human polypeptide	805	6	0.6	167	21	AA641298	Neisseria chimeric
733	6	0.6	150	18	AA623817	Cauliflower floral	806	6	0.6	167	22	AA626058	Novel human diago
734	6	0.6	150	19	AA669326	Brassica oleracea	807	6	0.6	168	21	AA616688	Arabidopsis thalia
735	6	0.6	150	19	AA643115	CAULIFLOWER gene p	808	6	0.6	168	22	AA620126	Novel human diago
736	6	0.6	150	19	AA643333	Cauliflower floral	809	6	0.6	168	22	AA603807	Rat partial mature
737	6	0.6	150	19	AA639137	Cauliflower floral	810	6	0.6	168	22	AAU19608	Human diagnostic a
738	6	0.6	150	21	AA619247	Amino acid sequenc	811	6	0.6	169	21	AA619335	Arabidopsis thalia
739	6	0.6	150	21	AA675556	B. oleracea var. b	812	6	0.6	169	21	AA630980	Arabidopsis thalia
740	6	0.6	150	21	AA678889	Cauliflower CAL am	813	6	0.6	169	21	AA633532	Arabidopsis thalia
741	6	0.6	150	22	AB623089	Novel human diago	814	6	0.6	169	21	AA642765	Arabidopsis thalia

815	6	0.6	169	22	AAE03810	Protein encoded by
816	6	0.6	170	22	AAE03811	Protein encoded by
817	6	0.6	170	22	AAE03811	Human ORF ORF47
818	6	0.6	170	21	AAE03811	Arabidopsis thalia
819	6	0.6	171	21	AAE03811	Protein involved i
820	6	0.6	171	21	AAE03811	Arabidopsis thalia
821	6	0.6	171	21	AAE03811	Arabidopsis thalia
822	6	0.6	171	22	AAE03811	Human haematologic
823	6	0.6	172	22	AAE03811	Proteinbacterium
824	6	0.6	172	22	AAE03811	Novel human diagno
825	6	0.6	172	22	AAE03811	Novel human diagno
826	6	0.6	173	22	AAE03811	Schizosaccharomyce
827	6	0.6	173	10	AAE03811	Sequence of a fusi
828	6	0.6	173	21	AAE03811	CD4 fragment bindi
829	6	0.6	173	22	AAE03811	Olfactory receptor
830	6	0.6	173	22	AAE03811	Human protein sequ
831	6	0.6	174	21	AAE03811	Zea mays protein f
832	6	0.6	174	22	AAE03811	Novel human diagno
833	6	0.6	174	22	AAE03811	Novel human diagno
834	6	0.6	175	22	AAE03811	Human protein sequ
835	6	0.6	175	22	AAE03811	Segment comprising
836	6	0.6	176	11	AAE03811	Sequence of chimae
837	6	0.6	176	22	AAE03811	Proteinbacterium
838	6	0.6	177	20	AAE03811	Human B-cell myeli
839	6	0.6	177	21	AAE03811	Arabidopsis thalia
840	6	0.6	179	21	AAE03811	Lung cancer associ
841	6	0.6	179	22	AAE03811	Proteinbacterium
842	6	0.6	179	22	AAE03811	Novel human diagno
843	6	0.6	181	18	AAE03811	Staphylococcus aur
844	6	0.6	181	21	AAE03811	Arabidopsis thalia
845	6	0.6	181	21	AAE03811	Arabidopsis thalia
846	6	0.6	181	21	AAE03811	Arabidopsis thalia
847	6	0.6	181	22	AAE03811	Novel signal trans
848	6	0.6	181	22	AAE03811	Novel human diagno
849	6	0.6	184	21	AAE03811	Human secreted pro
850	6	0.6	184	21	AAE03811	Human secreted pro
851	6	0.6	185	20	AAE03811	HIV-1 gp120 peptid
852	6	0.6	185	20	AAE03811	HIV gp120 protein
853	6	0.6	185	21	AAE03811	Arabidopsis thalia
854	6	0.6	185	21	AAE03811	Arabidopsis thalia
855	6	0.6	185	21	AAE03811	Sorangium cellulos
856	6	0.6	185	22	AAE03811	Escherichia coli p
857	6	0.6	185	22	AAE03811	Escherichia coli p
858	6	0.6	186	17	AAE03811	DNA mismatch repair
859	6	0.6	187	9	AAE03811	Sequence of yellow
860	6	0.6	187	19	AAE03811	Synthetic fish gro
861	6	0.6	187	21	AAE03811	Human ORF ORF173
862	6	0.6	187	22	AAE03811	Proteinbacterium
863	6	0.6	187	22	AAE03811	Novel human diagno
864	6	0.6	188	11	AAE03811	Fish growth hormon
865	6	0.6	188	13	AAE03811	Fish growth hormon
866	6	0.6	188	21	AAE03811	Arabidopsis thalia
867	6	0.6	188	21	AAE03811	Arabidopsis thalia
868	6	0.6	188	21	AAE03811	Arabidopsis thalia
869	6	0.6	188	22	AAE03811	Novel human diagno
870	6	0.6	188	22	AAE03811	Human novel extrac
871	6	0.6	189	20	AAE03811	Enterococcus faeca
872	6	0.6	189	20	AAE03811	Enterococcus faeca
873	6	0.6	190	20	AAE03811	Human B-cell myeli
874	6	0.6	190	21	AAE03811	Arabidopsis thalia
875	6	0.6	190	21	AAE03811	Arabidopsis thalia
876	6	0.6	190	22	AAE03811	Human NKp30 recept
877	6	0.6	191	12	AAE03811	Pre-pro AMF/CD4 fu
878	6	0.6	191	21	AAE03811	Arabidopsis thalia
879	6	0.6	191	21	AAE03811	Arabidopsis thalia
880	6	0.6	191	22	AAE03811	Drosophila melanog
881	6	0.6	191	22	AAE03811	Proteinbacterium
882	6	0.6	191	22	AAE03811	Rat interferon-lik
883	6	0.6	192	18	AAE03811	H. pylori cytoplas
884	6	0.6	192	22	AAE03811	Human polypeptide
885	6	0.6	193	21	AAE03811	Human ORF ORF262
886	6	0.6	193	22	AAE03811	Klebsiella pneumon
887	6	0.6	193	22	AAE03811	Novel human diagno
888	6	0.6	194	22	AAE03811	Novel human diagno
889	6	0.6	195	22	AAE03811	Olfactory receptor
890	6	0.6	195	22	AAE03811	Arabidopsis thalia
891	6	0.6	196	21	AAE03811	Arabidopsis thalia
892	6	0.6	196	21	AAE03811	Arabidopsis thalia
893	6	0.6	197	21	AAE03811	Amino acid sequenc
894	6	0.6	197	22	AAE03811	Human protein sequ
895	6	0.6	198	19	AAE03811	Mycothacterial hepa
896	6	0.6	198	19	AAE03811	Mycothacterial hepa
897	6	0.6	198	21	AAE03811	Arabidopsis thalia
898	6	0.6	198	22	AAE03811	Drosophila melanog
899	6	0.6	198	22	AAE03811	Human breast cance
900	6	0.6	199	20	AAE03811	Chlamydia pneumoni
901	6	0.6	199	22	AAE03811	Staphylococcus aur
902	6	0.6	199	22	AAE03811	Proteinbacterium
903	6	0.6	200	10	AAE03811	Fish growth hormon
904	6	0.6	200	16	AAE03811	Mouse B7-1 (IgV-1)
905	6	0.6	200	21	AAE03811	Arabidopsis thalia
906	6	0.6	200	21	AAE03811	Arabidopsis thalia
907	6	0.6	200	21	AAE03811	Arabidopsis thalia
908	6	0.6	201	20	AAE03811	Human B-cell myeli
909	6	0.6	201	21	AAE03811	Arabidopsis thalia
910	6	0.6	201	21	AAE03811	Arabidopsis thalia
911	6	0.6	201	22	AAE03811	Human protein sequ
912	6	0.6	203	10	AAE03811	Fish growth hormon
913	6	0.6	203	11	AAE03811	Fish growth hormon
914	6	0.6	203	16	AAE03811	CD4 domains D1-D2
915	6	0.6	203	17	AAE03811	CD4 D1-D2 domains
916	6	0.6	203	21	AAE03811	Novel human diagno
917	6	0.6	203	21	AAE03811	Novel human diagno
918	6	0.6	203	22	AAE03811	Novel human diagno
919	6	0.6	203	22	AAE03811	Novel human diagno
920	6	0.6	204	10	AAE03811	Fish growth hormon
921	6	0.6	204	12	AAE03811	Fish growth hormon
922	6	0.6	204	18	AAE03811	Yellowtail tuna gr
923	6	0.6	204	18	AAE03811	Yellow tail/fin tu
924	6	0.6	204	20	AAE03811	Acetyl transferase
925	6	0.6	204	22	AAE03811	Drosophila melanog
926	6	0.6	204	22	AAE03811	Proteinbacterium
927	6	0.6	204	22	AAE03811	Human transmembran
928	6	0.6	204	22	AAE03811	S. epidermidis ope
929	6	0.6	204	22	AAE03811	S. epidermidis ope
930	6	0.6	204	22	AAE03811	MS-15 lissue-pref
931	6	0.6	205	19	AAE03811	Human polypeptide
932	6	0.6	205	22	AAE03811	Malassezia fungus
933	6	0.6	206	18	AAE03811	Drosophila melanog
934	6	0.6	206	22	AAE03811	Arabidopsis thalia
935	6	0.6	207	21	AAE03811	Arabidopsis thalia
936	6	0.6	208	21	AAE03811	Arabidopsis thalia
937	6	0.6	208	21	AAE03811	Arabidopsis thalia
938	6	0.6	208	21	AAE03811	Arabidopsis thalia
939	6	0.6	208	22	AAE03811	Novel human diagno
940	6	0.6	208	22	AAE03811	S. aureus conserved
941	6	0.6	209	14	AAE03811	Prod. of the Acet
942	6	0.6	209	19	AAE03811	Human COP9 protein
943	6	0.6	209	21	AAE03811	Arabidopsis thalia
944	6	0.6	209	21	AAE03811	Arabidopsis thalia
945	6	0.6	209	21	AAE03811	Arabidopsis thalia
946	6	0.6	209	21	AAE03811	Arabidopsis thalia
947	6	0.6	210	21	AAE03811	Arabidopsis thalia
948	6	0.6	210	21	AAE03811	Arabidopsis thalia
949	6	0.6	210	21	AAE03811	Arabidopsis thalia
950	6	0.6	210	22	AAE03811	Arabidopsis thalia
951	6	0.6	210	22	AAE03811	Proteinbacterium
952	6	0.6	211	15	AAE03811	Human G protein-co
953	6	0.6	211	15	AAE03811	Hepatitis type-c v
954	6	0.6	211	21	AAE03811	H. pylori ORF 14ce
955	6	0.6	211	21	AAE03811	Arabidopsis thalia
956	6	0.6	211	21	AAE03811	Arabidopsis thalia
957	6	0.6	211	21	AAE03811	Arabidopsis thalia
958	6	0.6	211	21	AAE03811	Arabidopsis thalia
959	6	0.6	212	20	AAE03811	Protein encoded by
960	6	0.6	212	20	AAE03811	Amino acid sequenc

```
961 6 0.6 212 21 AAB27945 Human secreted pro
962 6 0.6 212 22 AAU37356 Staphylococcus aur
963 6 0.6 212 22 AAU37504 Staphylococcus aur
964 6 0.6 212 22 AAB95582 Human protein sequ
965 6 0.6 212 22 AAB36585 Human FLEXMT-7 pro
966 6 0.6 213 19 AAY11072 H. pylori ORF hp2p
967 6 0.6 213 13 AAG47012 Arabidopsis thalia
968 6 0.6 214 16 AAR82901 Mouse B7-1 (Igv-11
969 6 0.6 214 21 AAG20979 Arabidopsis thalia
970 6 0.6 216 20 AAY33537 A. thaliana CBF3 P
971 6 0.6 216 21 AAG44176 Arabidopsis thalia
972 6 0.6 216 21 AAB03745 Dehydration respon
973 6 0.6 216 21 AAB03753 Dehydration respon
974 6 0.6 216 21 AAY82475 A. thaliana transc
975 6 0.6 216 22 AAM25866 Human protein sequ
976 6 0.6 216 22 AAE02569 A. thaliana transc
977 6 0.6 217 12 AAR15150 PC04-gelsolin sequ
978 6 0.6 218 22 AAM93195 Human polypeptide,
979 6 0.6 218 22 AAG81734 S. epidermidis ope
980 6 0.6 219 19 AAM80693 S. pneumoniae DNA
981 6 0.6 220 19 AAM79323 Staphylococcus aur
982 6 0.6 220 21 AAY35523 Chlamydia pneumoni
983 6 0.6 220 21 AAB38499 Fragment of human
984 6 0.6 220 21 AAG16490 Arabidopsis thalia
985 6 0.6 220 21 AAG37994 Arabidopsis thalia
986 6 0.6 221 20 AAY36863 Protein which is s
987 6 0.6 221 21 AAG21867 Arabidopsis thalia
988 6 0.6 222 21 AAY51400 S. pombe Weel cata
989 6 0.6 222 21 AAB62518 S. pombe Weel prot
990 6 0.6 223 13 AAR29923 SAP. Homo sapiens
991 6 0.6 223 13 AAG06620 Arabidopsis thalia
992 6 0.6 223 21 AAG14149 Arabidopsis thalia
993 6 0.6 223 21 AAG38253 Arabidopsis thalia
994 6 0.6 223 22 AAG05904 Novel human diagno
995 6 0.6 224 13 AAR25675 Monoclonal antibod
996 6 0.6 224 18 AAM14788 FRBP-LCK-SH2 fusio
997 6 0.6 224 20 AAM96823 A fusion protein o
998 6 0.6 224 21 AAY94892 Human protein clon
999 6 0.6 224 22 AAG15369 Human human diagno
1000 6 0.6 224 22 AAE02549 A. thaliana transc
```

ALIGNMENTS

```
RESULT 1
ID AAM79303 standard; Protein; 502 AA.
XX
AC AAM79303;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human protein SEQ ID NO 1965.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue; growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation.
XX
OS Homo sapiens.
XX
PN WO200157190-A2.
XX
PD 09-AUG-2001.
XX
PE 05-FEB-2001; 2001WO-US04098.
XX
PR 03-FEB-2000; 2000US-0496914.
PR 27-APR-2000; 2000US-0560875.
PR 20-JUN-2000; 2000US-0598075.
PR 19-JUL-2000; 2000US-0620325.
PR 01-SEP-2000; 2000US-0654936.
```

```
PR 15-SEP-2000; 2000US-0663561.
PR 20-OCT-2000; 2000US-0693325.
PR 30-NOV-2000; 2000US-0728422.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y,
PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AJ, Yang Y, Wejberman T, Goodrich R;
XX
DR WPI: 2001-476283/51.
DR N-PSDB: AAK52436.
XX
PT Nucleic acids encoding polypeptides with cytokine-like activities,
PT useful in diagnosis and gene therapy -
XX
PS Claim 20; Page 4376-4377; 6221pp; English.
XX
CC The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC encoded polypeptides (AAM78323-AAK80302) that exhibit actively elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activity/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.
CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
CC (AAM80020) are omitted as the relevant pages from the sequence listing
CC were missing at the time of publication.
XX
SQ Sequence 502 AA;
XX
Query Match 34.2%; Score 367; DB 22; Length 502;
Best Local Similarity 100.0%; Pred. NO. 0;
Matches 367; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 708 MCRINDENAMNELARACLIHMEVEFAIRVRRIGNVGIVMSLEQIKGIEDYNLAGHIA 767
DB 1 mcrllndeawmelaracchhmevefaiirvrrignvgivmsleqikgedynllagha 60
QY 768 MFTNDYNLAQDLYLASCPALAEKMRDLQHWDSALQAKHLAPQPIFISKEVIAIQLEF 827
DB 61 mftndynlaqdllylascpialaemrldqhwdsalqlakhlapqpifiskeyaqllef 120
QY 828 AGDYVNATLAHYEKGTGDKKEHDEACLAGVAQMSIRMGDIRGYNQAKHPRVLRKRCG 887
DB 121 agdyvnalalyekgltgdhkehdeacclagvaqmslrmgdirgvnqalkhpsrvlkrdcg 180
QY 888 ALLENMKOFSEAAQLYEKGLYDYDKAASYIRSKNNAKVGDLLPHVSPSRHLYQYAKKEA 947
DB 181 allenmkofseaaqlyeykglydydkaasyirsknnakvgdllphvssprhlyqyakkea 240
QY 948 DGRYKFAVVAYENAKQWOSVIRIYLDHLNPNPKAVNIYRETQSLDCAKWAFAFLQGDY 1007
DB 241 dgrykfaavayenakwqsviriyldhlnpnpekavniyretqslgdkwafafflqgdy 300
QY 1008 GSAIGFLVMSKCNNEAFITLAQOHNKMETYADITIGSEDTNEDYOSIALYFEGERKYLAG 1067
DB 301 gsaigflvmskcnneafitlaqohnmkemyaditigsedtnedysialyfegekrylag 360
QY 1068 KFFLLCG 1074
DB 361 kffllcg 367
XX
RESULT 2
ID AAM80287 standard; Protein; 515 AA.
XX
```


AC AAM80287;
 XX
 DT 06-NOV-2001 (first entry)
 XX
 DE Human protein SEQ ID NO 3933.
 XX
 XX Human: cytokine; cell proliferation; cell differentiation; gene therapy;
 KM vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KM tissue growth factor; immunomodulatory; cancer; leukaemia;
 KM nervous system disorder; arthritis; inflammation.
 XX
 OS Homo sapiens.
 XX
 PN WO200157190-A2.
 XX
 PD 09-AUG-2001.
 XX
 PE 05-FEB-2001; 2001WO-US04098.
 XX
 PR 03-FEB-2000; 2000US-0496914.
 PR 27-APR-2000; 2000US-0560875.
 PR 20-JUN-2000; 2000US-0598075.
 PR 19-JUL-2000; 2000US-0620325.
 PR 01-SEP-2000; 2000US-0654936.
 PR 15-SEP-2000; 2000US-0663561.
 PR 20-OCT-2000; 2000US-0693325.
 PR 30-NOV-2000; 2000US-0728422.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
 PI Zhao QA, Wang D, Wang J, Ren F, Chen R, Wang ZW;
 PI Xue AJ, Yang Y, Wejhrman T, Goodrich R.
 XX
 DR WPI; 2001-476283/51.
 XX
 DR N-PSDB; AAK53420.
 XX
 PT Nucleic acids encoding polypeptides with cytokine-like activities,
 PT useful in diagnosis and gene therapy -
 XX
 PS Claim 20: Page 471; 6221pp; English.
 XX
 XX The invention relates to polynucleotides (AAK51456-AAK53435) and the
 CC encoded polypeptides (AAM80287-AAK80302) that exhibit activity relating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activity/inhibitor activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation.
 CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
 CC (AAM80020) are omitted as the relevant pages from the sequence listing
 CC were missing at the time of publication.
 CC
 XX Sequence 515 AA:
 SO

Query Match 25.8%; Score 277; DB 22; Length 515;
 Best Local Similarity 100.0%; Pred. No. 1.5e-280;
 Matches 277; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 798 HNSAQLAHRLAPDQPTFSKEAIOLEFAGDYVNAHLAYEKGITDDNKEHDEACLAGV 857
 DB 7 hwsaqlahrlapdqptfiskayqlafagdyvnalahyexyltgdnkheacagv 66
 QY 858 AOKSIRKGDTRRGVNAOLKTPSRVLRKDCGALLENMKOFSEAAQLYKGLYDOKAAASYVI 917
 DB 67 agmsirgdttrrgvnaolktpsrvlrkdcgallennkfseaaqlkylgldkaasyvi 126
 QY 918 RSKNMAKVGDLLPHVSSPKITHLOYAKAKKADGRYKEAVVAENAKOMQSVIRIYDLHLNN 977

DB 127 rsknmakvgdllphvsspkithloyakakkeadgyrkeavvayenakqgsviriydlhlmm 186
 QY 978 PEKAVNIVRETQSLDGAKMKARREFLOLDGYSALQFLVMSKCNNEAFTLAQGNHKEITYA 1037
 DB 187 pekavnivretqslgdakmkarrefflqldgsalqflvmskcnneaftlaqgnhmeiya 246
 QY 1038 DIIGSEDTTNEEDYOSIALYPEGKRYLQAGKFFLLCG 1074
 DB 247 diigsedttneedyosalylegekrlylqagkffllcg 283

RESULT 3
 AAU14751
 ID AAU14751 standard; Protein; 138 AA.
 AC AAU14751;
 DT 24-OCT-2001 (first entry)
 DE Novel bone marrow polypeptide #150.
 XX
 XX Bone marrow; diagnostic; therapeutic; gene therapy; antigenic;
 KM haematopoiesis; myeloid; lymph cell disorder; tissue regeneration;
 KM wound healing; nutritional supplement; immune disorder;
 KM severe combined immunodeficiency; SCID.
 XX
 OS Homo sapiens.
 XX
 PN WO200157187-A2.
 XX
 PD 09-AUG-2001.
 XX
 PE 05-FEB-2001; 2001WO-US03782.
 XX
 PR 03-FEB-2000; 2000US-0496914.
 PR 20-JUN-2000; 2000US-0598075.
 PR 19-JUL-2000; 2000US-0620325.
 PR 30-NOV-2000; 2000US-0250683.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Ford JE, Boyle BJ, Tang YT, Liu C, Asundi V, Zhou P, Xue AJ;
 PI Ren F, Drmanac RT;
 XX
 DR WPI; 2001-488875/53.
 XX
 DR N-PSDB; AAS23056.
 XX
 PT Nucleic acids encoding bone marrow polypeptides, useful in diagnostic
 PT and gene therapy -
 XX
 PS Claim 10: Page 135; 392pp; English.
 XX
 XX AAU14602-AAU14794 represent novel bone marrow polypeptides of the
 CC invention. The proteins and corresponding coding sequences may be used
 CC in the prevention, diagnosis and treatment of diseases associated with
 CC inappropriate bone marrow polypeptide expression. For example, to treat
 CC disorders associated with decreased expression by rectifying mutations
 CC or deletions in a patient's genome that affect the activity of the
 CC polypeptides by expressing inactive proteins or to supplement the
 CC patient's own production of the polypeptides. Additionally, the nucleic
 CC acids may be used to produce the polypeptides, by inserting the nucleic
 CC acids into a host cell and culturing the cell to express the protein.
 CC The nucleic acid and its complementary sequences may also be used as DNA
 CC probes in diagnostic assays to detect and quantitate the presence of
 CC similar nucleic acid sequences in samples, and therefore which patients
 CC may be in need of restorative therapy. The proteins may also be used as
 CC antigens in the production of antibodies against bone marrow proteins
 CC and in assays to identify modulators of their expression and activity.
 CC The anti-bone marrow protein antibodies and antigens may also be used
 CC to down regulate expression and activity. The antibodies may also be used
 CC as diagnostic agents for detecting the presence of the protein in samples
 CC (e.g. by enzyme linked immunosorbant assay (ELISA)). The proteins

CC may be used to regulate haematopoiesis activity, and consequently in the
CC treatment of myeloid or lymph cell disorders; in tissue regeneration,
CC such as wound healing; as a nutritional supplement; and in treatment of
CC immune disorders such as severe combined immunodeficiency (SCID).
XX
SQ Sequence 138 AA;

Query Match 7.4%; Score 80; DB 22; Length 138;
Best Local Similarity 100.0%; Pred. No. 6,3e-75;
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 223 NQFLLMKDDRTSAESHSISVYLKKTLPFLNLNPPNADLEPQDFGNVICWICWD 282
DB 59 nmqfilmkmdrttsaesmslv1gkklflfnlnepdnadelqgdfgnlvcywygd 118
|||||

OY 283 GRIMIGSCGHFVISTHNG 302
DB 119 grlmigfscghfvisthng 138
|||||

RESULT 4
AU14657
ID AU14657 standard; Protein; 244 AA.
XX
AC AU14657;
XX
DT 24-OCT-2001 (first entry)
XX

DE Novel bone marrow polypeptide #56.

KW Bone marrow; diagnostic; therapeutic; gene therapy; antigenic;
KW haematopoiesis; myeloid; lymph cell disorder; tissue regeneration;
KW wound healing; nutritional supplement; immune disorder;
KW severe combined immunodeficiency; SCID.

OS Homo sapiens.

PN WO200157187-A2.

PD 09-AUG-2001.

PE 05-FEB-2001; 2001WO-US03782.

XX 03-FEB-2000; 2000US-0496914.

PR 20-JUN-2000; 2000US-0598075.

PR 19-JUL-2000; 2000US-0620325.

PR 30-NOV-2000; 2000US-0250683.

XX (HYSE-) HYSEQ INC.

XX Ford JE, Boyle BJ, Tang YT, Liu C, Asundi V, Zhou P, Xue AJ;
PI Ren F, Drmanac RT;

XX WPI: 2001-488875/53.

DR N-PSDB; AAS22962.

XX Nucleic acids encoding bone marrow polypeptides, useful in diagnostic
PT and gene therapy -
PT
XX

PS Claim 10; Page 239-240; 392pp; English.

XX AU14602-AU14794 represent novel bone marrow polypeptides of the
CC invention. The proteins and corresponding coding sequences may be used
CC in the prevention, diagnosis and treatment of diseases associated with
CC inappropriate bone marrow polypeptide expression. For example, to treat
CC disorders associated with decreased expression by rectifying mutations
CC or deletions in a patient's genome that affect the activity of the
CC polypeptides by expressing inactive proteins or to supplement the
CC patient's own production of the polypeptide. Additionally, the nucleic
CC acids may be used to produce the polypeptides, by inserting the nucleic
CC acids into a host cell and culturing the cell to express the protein.
CC The nucleic acid and its complementary sequences may also be used as DNA

CC probes in diagnostic assays to detect and quantitate the presence of
CC similar nucleic acid sequences in samples, and therefore which patients
CC may be in need of restorative therapy. The proteins may also be used as
CC antigens in the production of antibodies against bone marrow proteins
CC and in assays to identify modulators of their expression and activity
CC The anti-bone marrow protein antibodies and antagonists may also be used
CC to down regulate expression and activity. The antibodies may also be used
CC as diagnostic agents for detecting the presence of the protein in samples
CC (e.g. by enzyme linked immunosorbent assay (ELISA)). The proteins
CC may be used to regulate haematopoiesis activity, and consequently in the
CC treatment of myeloid or lymph cell disorders; in tissue regeneration, of
CC such as wound healing; as a nutritional supplement; and in treatment of
CC immune disorders such as severe combined immunodeficiency (SCID).
XX
SQ Sequence 244 AA;

Query Match 5.8%; Score 62; DB 22; Length 244;
Best Local Similarity 100.0%; Pred. No. 7,3e-36;
Matches 62; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 100 MDMKDGVLAVIAKSSCIYLMDFNTKTSQDNGMRDMSFLMSKVSFLAVGTGK 159
DB 1 mdwkdgdv1aviaekssc1yldanfntktsqldngmrdmsflmskvsflavgtvg 60
|||||

OY 160 NL 161
DB 61 nl 62
||

RESULT 5
AU33013
ID AU33013 standard; Protein; 72 AA.
XX
AC AU33013;
XX
DT 18-DEC-2001 (first entry)
XX

DE Novel human secreted protein #3504.

KW Human; vaccination; gene therapy; nutritional supplement;
KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
KW immune suppression; immune stimulation; anti-inflammatory; leukemia.

OS Homo sapiens.

PN WO200179449-A2.

PD 25-OCT-2001.

PE 16-APR-2001; 2001WO-US08656.

XX 18-APR-2000; 2000US-0552929.

PR 26-JAN-2001; 2001US-0770160.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RT;

DR WPI: 2001-611725/70.

XX Nucleic acids encoding a range of human polypeptides, useful in genetic
PT vaccination, testing and therapy -
PT
XX

PS Claim 20; Page 698; 765pp; English.

XX The invention relates to novel human secreted polypeptides. The
CC polypeptides and antibodies to the polypeptides are useful for
CC determining the presence of or predisposition to a disease associated
CC with altered levels of polypeptide. The polypeptides are also useful for
CC identifying agents (agonists and antagonists) that bind to them. Cells
CC expressing the proteins are useful for identifying a therapeutic agent
CC for use in treatment of a pathology related to aberrant expression or

CC physiological interactions of the polypeptide. Vectors comprising
 CC the nucleic acids encoding the polypeptides and cells genetically
 CC engineered to express them are also useful for producing the proteins.
 CC The proteins are useful in genetic vaccination, testing and
 CC therapy, and can be used as nutritional supplements. They may be used to
 CC increase stem cell proliferation; to regulate haematopoiesis; and in
 CC bone, cartilage, tendon and/or nerve tissue growth or regeneration;
 CC immune suppression and/or stimulation; as anti-inflammatory agents; and
 CC in treatment of leukaemias. AAU29510-AAU33304 represent the amino acid
 CC sequences of novel human secreted proteins of the invention.

XX Sequence 72 AA;

Query Match 3.0%; Score 32; DB 22; Length 72;
 Best Local Similarity 100.0%; Pred. No. 5.8e-25;
 Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 295 VVISTHTGELGDEIFQARNHKDNLTSIAVSQT 326
 ||||||||||||||||||||||||||||||||
 Db 29 vvishtgtelgdeifqarnhknltsiavsqt 60

RESULT 6

AAU33123
 ID AAU33123 standard; Protein: 237 AA.

XX AAU33123;

AC 03-SEP-2001 (first entry)

DE Human colon cancer antigen protein SEQ ID NO:5475.

KW Human: colon cancer; colon cancer antigen; diagnosis; detection;
 KM colorectal carcinoma; chromosome 1.

XX Homo sapiens.

PN WO200122920-A2.

PD 05-APR-2001.

PF 28-SEP-2000; 2000WO-US26524.

PR 29-SEP-1999; 99US-0157137.

PR 03-NOV-1999; 99US-0163280.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Ruben SM, Barash SC, Birse CE, Rosen CA;

PT WPI; 2001-235357/24.

DR N-PSDB; AAH34116.

XX Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
 PT useful for preventing, diagnosing and/or treating colorectal cancers -

XX Claim 11: Page 7076-7077; 9803pp; English.

XX AAH32943 to AAH37195 and AAU33123 to AAU33123 represent human colon
 CC cancer-associated nucleic acid molecules (N) and proteins (P), where
 CC the proteins are collectively known as colon cancer antigens. The colon
 CC cancer antigens have cytostatic activity and can be used in gene
 CC therapy and vaccine production. N and P may be used in the prevention,
 CC diagnosis and treatment of diseases associated with inappropriate P
 CC expression. For example, N and P may be used to treat disorders
 CC associated with decreased expression by rectifying mutations or deletions
 CC in a patient's genome that affect the activity of P by expressing
 CC inactive proteins or to supplement the patients own production of P.
 CC Additionally, N may be used to produce the colon cancer-associated PS,
 CC by inserting the nucleic acids into a host cell and culturing the cell
 CC to express the proteins. N and P can be used in the prevention, diagnosis
 CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204

CC and AAU33123 represent sequences used in the exemplification of the
 CC present invention.
 CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were
 CC missing at time of publication, meaning no sequences are present for
 CC SEQ ID NO:1027 to 1052, 7921 and 7922.

XX Sequence 237 AA;

Query Match 2.3%; Score 25; DB 22; Length 237;
 Best Local Similarity 100.0%; Pred. No. 3.7e-17;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1050 YOSIALYFEGEKRYIQAGKFFELCG 1074
 ||||||||||||||||||||||||||||||||
 Db 1 ysiailyfegekryiqagkffilcg 25

RESULT 7

AAU33123
 ID AAU33123 standard; Protein: 109 AA.

XX AAU33123;

AC 18-DEC-2001 (first entry)

DE Novel human secreted protein #3614.

XX Human: vaccination; gene therapy; nutritional supplement;
 KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
 KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.

XX Homo sapiens.

PN WO200179449-A2.

PD 25-OCT-2001.

PF 16-APR-2001; 2001WO-US06656.

PR 18-APR-2000; 2000US-0552929.

PR 26-JAN-2001; 2001US-0770160.

PA (HYSE-) HYSEQ INC.

PI Tang YT, Liu C, Dmanac RT;

PT WPI; 2001-611725/70.

DR Nucleic acids encoding a range of human polypeptides, useful in genetic
 PT vaccination, testing and therapy -

XX Claim 20: Page 710; 765pp; English.

XX The invention relates to novel human secreted polypeptides. The
 CC polypeptides and antibodies to the polypeptides are useful for
 CC determining the presence of or predisposition to a disease associated
 CC with altered levels of polypeptide. The polypeptides are also useful for
 CC identifying agents (agonists and antagonists) that bind to them. Cells
 CC expressing the proteins are useful for identifying a therapeutic agent
 CC for use in treatment of a pathology related to aberrant expression or
 CC physiological interactions of the polypeptide. Vectors comprising
 CC the nucleic acids encoding the polypeptides and cells genetically
 CC engineered to express them are also useful for producing the proteins.
 CC The proteins are useful in genetic vaccination, testing and
 CC therapy, and can be used as nutritional supplements. They may be used to
 CC increase stem cell proliferation; to regulate haematopoiesis; and in
 CC bone, cartilage, tendon and/or nerve tissue growth or regeneration;
 CC immune suppression and/or stimulation; as anti-inflammatory agents; and
 CC in treatment of leukaemias. AAU29510-AAU33304 represent the amino acid
 CC sequences of novel human secreted proteins of the invention.

XX Sequence 109 AA;

Query Match 1.8%; Score 19; DB 22; Length 109;
 Best Local Similarity 100.0%; Pred. No. 3.4e-11;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1032 KMEIYADIIIGSEDTTNEYD 1050
 ||||||||||||||||
 DB 72 KMEIYADIIIGSEDTTNEYD 90

RESULT 8
 AAY41351
 ID AAY41351 standard; Protein: 55 AA.
 XX
 AC AAY41351;
 XX
 DT 02-DEC-1999 (first entry)
 XX
 DE Human secreted protein encoded by gene 44 clone HIXEY51.
 XX
 KW Human; secreted protein; fusion protein; gene therapy; protein therapy;
 KW diagnosis; tissue; cancer; tumor; neurodegenerative disorder; leukaemia;
 KW developmental abnormality; foetal deficiency; blood; allergy; renal;
 KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
 KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
 KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
 KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
 KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
 XX
 OS Homo sapiens.
 XX
 PN WO9947540-A1.
 XX
 PD 23-SEP-1999.
 XX
 PE 18-MAR-1999; 99WO-US05804.
 XX
 PR 19-MAR-1998; 98US-0078563.
 PR 19-MAR-1998; 98US-0078566.
 PR 19-MAR-1998; 98US-0078573.
 PR 19-MAR-1998; 98US-0078574.
 PR 19-MAR-1998; 98US-0078576.
 PR 19-MAR-1998; 98US-0078577.
 PR 19-MAR-1998; 98US-0078578.
 PR 19-MAR-1998; 98US-0078579.
 PR 19-MAR-1998; 98US-0078581.
 PR 01-APR-1998; 98US-0080312.
 PR 01-APR-1998; 98US-0080313.
 PR 01-APR-1998; 98US-0080314.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Ruben SM, Ni J, Rosen CA, Yu G, Young PE, Feng P, Soppet DR;
 PI Wel Y, Endress GA, Duan RD, Kyaw H, Ebner R, Lafleur DW;
 PI Olsen HS, Shi Y, Moore PA;
 XX
 DR WPI: 1999-562050/47.
 DR N-PSDB: AA224854.
 XX
 PT New isolated human genes, useful for diagnosis and treatment of e.g.
 PT cancers, neurological disorders, immune diseases, inflammation or blood
 PT disorders -
 XX
 PS Claim 11; Page 385; 484pp; English.
 XX
 CC This sequence represents a secreted human protein encoded by the gene
 CC clone detailed in the descriptor line. The gene can be used to generate
 CC fusion proteins by linking to the gene to a human immunoglobulin Fc
 CC portion (e.g. AA224802) for increasing the stability of the fused
 CC protein as compared to the human protein only.
 CC The invention relates to 95 novel genes and their fragments (nucleic
 CC acid sequences: AA224811-224907; amino acid sequences AAY41308-Y41404)

CC which are useful for preventing, treating or ameliorating medical
 CC conditions e.g. by protein or gene therapy. Also, pathological
 CC conditions can be diagnosed by determining the amount of the new
 CC polypeptides in a sample or by determining the presence of mutations in
 CC the new polynucleotides. Specific uses are described for each of the 95
 CC polynucleotides, based on which tissues they are most highly expressed in
 CC (see AA224811 for described uses).
 CC
 XX
 SQ Sequence 55 AA;

Query Match 0.7%; Score 7; DB 20; Length 55;
 Best Local Similarity 100.0%; Pred. No. 66;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 873 QALKHPS 879
 |||||||
 DB 37 qalkhps 43

RESULT 9
 AAR38501
 ID AAR38501 standard; protein: 57 AA.
 XX
 AC AAR38501;
 XX
 DT 28-OCT-1993 (first entry)
 XX
 DE M. bovis pili protein submolecular unit.
 XX
 KW Antibodies; whole pili binding; basis; vaccine; bacterial infection;
 KW ruminant footrot infection; sheep; type IV pillated bacteria.
 XX
 OS Moraxella bovis.
 XX
 PN WO9311791-A.
 XX
 PD 24-JUN-1993.
 XX
 PF 17-DEC-1992; 92WO-US11085.
 XX
 PR 18-DEC-1991; 91US-0809762.
 XX
 PA (UYOR-) UNIV OREGON HEALTH SCI.
 XX
 PI Smith AW;
 XX
 DR WPI: 1993-213824/26.
 XX
 PT Antigenic preparation - stimulates production of antibodies binding
 PT to pili protein of type IV pillated bacteria, useful in
 PT vaccine compsn.
 XX
 PS Claim 12; Page 26; 44pp; English.
 XX
 CC The sequence is that of a submolecular unit of Moraxella bovis
 CC pili protein which corresponds to at least one epitope common to
 CC structural pili proteins of Type IV pillated bacteria. It is
 CC capable of eliciting antibodies (Abs) which bind to whole pili of
 CC type IV bacteria. The ability of the submolecular unit to produce
 CC Abs which bind to whole pili provides the basis for vaccines against
 CC type IV bacterial infections, e.g. footrot infection in ruminants.
 CC
 XX
 SQ Sequence 57 AA;
 QY 481 AALFEKG 487
 |||||||
 DB 49 aalfeqk 55

RESULT 10
 AAB48729
 ID AAB48729 standard; Protein; 58 AA.
 AC AAB48729;
 DT 09-MAR-2001 (first entry)
 DE Mouse liver growth hormone-induced clone 5 ORF #6.
 XX
 XX
 KW Mouse: growth hormone; GH regulatable gene; liver pathology; hypertrophy;
 KW hepatocellular lesion; hyperplasia; altered expression level; clone 5;
 KW diagnostic marker; gigantism; acromegaly; diabetes; hepatotropic;
 KW transgenic animal; drug screening; drug discovery; murine;
 KW open reading frame; ORF.
 XX
 XX
 OS Mus sp.
 XX
 XX
 PN MO200066787-A2.
 XX
 PD 09-NOV-2000.
 XX
 XX
 PF 05-MAY-2000; 2000MO-US12366.
 XX
 PR 05-MAY-1999; 99US-0132663.
 XX
 XX
 PA (UYOH-) UNIV OHIO.
 XX
 PI Kopschick JJ, Tjong J;
 XX
 DR WPI: 2001-007239/01.
 DR N-PSDB; AAC87257.
 XX
 PT Diagnosing abnormal levels of growth hormone activity in liver
 PT comprising assaying growth transcriptional activity and protein
 PT expression level of hormone-regulatable liver genes, as diagnostic
 PT markers of liver pathology -
 XX
 XX
 PS Example 2; Page 53; 65pp; English.
 XX
 CC The invention relates to a method of diagnosing abnormal levels of
 CC growth hormone (GH) activity in the liver, or predicting a change in the
 CC condition of the liver in response to abnormal GH activity. The method
 CC involves correlating the level of expression of certain specific genes
 CC with the level of GH activity in the liver, or with an expected change
 CC in the condition of the liver as the result of GH activity. Excessive GH
 CC activity in the liver is thought to be deleterious to health, causing an
 CC increase in liver size as a consequence of both hyperplasia and
 CC hepatocyte hypertrophy, and hepatocellular lesions which progress with
 CC age. Studies in transgenic mice which express high levels of bovine
 CC growth hormone identified a number of genes whose expression in the
 CC liver is altered by high GH levels. The genes which are upregulated are
 CC those encoding alpha-fetoprotein, corticosteroid binding globulin,
 CC fetuin, ras8-interacting protein, paroxonase-3, cytochrome P450I1A,
 CC 5-2 hydroxyacid oxidase, interferon alpha/beta receptor, growth hormone
 CC receptor, proteasome z-subunit, and coagulation factor V. Two novel
 CC genes, clone 5 (AAC87257) and clone 45 (AAC87258) are also upregulated
 CC in response to abnormally high GH levels. Conversely, expression of the
 CC gene encoding z-beta-hydroxysteroid dehydrogenase/delta-5-delta-4
 CC isomerase appears to be downregulated. The invention also relates to
 CC transgenic nonhuman mammals comprising a GH-induced transgene which
 CC exhibit or have a propensity to develop a liver pathology; an assay for
 CC drugs which inhibit the development of, or which treat a liver pathology,
 CC comprising administering the drug to the transgenic animal; and
 CC preventing or treating a liver pathology in a patient comprising
 CC administering a drug which inhibits the expression of a GH-regulatable
 CC gene. The method of the invention is used for diagnosing abnormal levels of GH
 CC activity in the liver or predicting a change in the condition of the
 CC liver in response to abnormal levels of GH activity. The GH-regulatable
 CC liver genes and proteins are useful as diagnostic markers of liver
 CC pathology. Assays for the expression of these genes is useful for the

CC diagnosis of liver pathologies associated with gigantism or acromegaly or
 CC with diabetes, as other causative agents may act directly or indirectly
 CC upon the same genes. Sequence AAB48724-BA8739 represent the open reading
 CC frames (ORFs) of the novel mouse liver cDNA clone 5.
 XX
 XX
 SO Sequence 58 AA;
 XX
 Query Match 0.7%; Score 7; DB 22; Length 58;
 Best Local Similarity 100.0%; Pred. No. 69;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 874 ALKPSR 880
 |||||
 Db 42 alkhpsr 48
 XX
 RESULT 11
 AAG03014
 ID AAG03014 standard; Protein; 61 AA.
 XX
 AC AAG03014;
 XX
 DT 06-OCT-2000 (first entry)
 XX
 DE Human secreted protein, SEQ ID NO: 7095.
 XX
 KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
 KW gene therapy; chromosome mapping.
 XX
 OS Homo sapiens.
 XX
 PN EP1033401-A2.
 XX
 PD 06-SEP-2000.
 XX
 PF 21-FEB-2000; 2000EP-0200610.
 XX
 PR 26-FEB-1999; 99US-0122487.
 XX
 XX
 PA (GEST) GENSET.
 XX
 PI Dumas Milne Edwards J, Duclert A, Giordano J;
 XX
 DR WPI: 2000-500381/45.
 DR N-PSDB; AAC03020.
 XX
 XX
 PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
 PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
 XX
 XX
 PS Claim 13; SEQ ID 7095; 71pp + CD-ROM; English.
 XX
 CC The present sequence is a polypeptide encoded by one of a large number
 CC of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs
 CC were prepared from total human RNAs or polyA+ RNAs derived from 30
 CC different tissues. EST sequences usually correspond mainly to the 3'
 CC untranslated region (UTR) of the mRNA because they are often obtained
 CC from oligo-dT primed cDNA libraries. Such ESTs are not well suited for
 CC isolating cDNA sequences derived from the 5' ends of mRNAs and even in
 CC those cases where longer cDNA sequences have been obtained, the full 5'
 CC UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'
 CC ends and can therefore be used to obtain full length cDNAs with genomic
 CC DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and
 CC chromosome mapping procedures. They are used to obtain upstream
 CC regulatory sequences and to design expression and secretion vectors.
 XX
 SO Sequence 61 AA;
 Query Match 0.7%; Score 7; DB 21; Length 61;
 Best Local Similarity 100.0%; Pred. No. 72;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 986 RETOSUD 992
 11111111
 DB 28 retgsid 34

RESULT 12

ID AAB39306 standard; Protein; 62 AA.

AC AAB39306;

DT 02-FEB-2001 (first entry)

DE Human secreted protein sequence encoded by gene 47 SEQ ID NO:186.

KM Human; secreted protein; immunosuppressive; antiarthritic; antirheumatic;
 KM antiproliferative; cytostatic; cardiant; vasotropic; cerebroprotective;
 KM neotropic; neuroprotective; antibacterial; virucide; fungicide; neoplasm;
 KM ophthalmological; autoimmune disease; rheumatoid arthritis; angiogenesis;
 KM hyperproliferative disorder; cardiovascular disorder; infection;
 KM cerebrovascular disorder; nervous system disorder; ocular disorder;
 KM wound healing; chemotaxis.

OS Homo sapiens.

PN WO200056754-A1.

PD 28-SEP-2000.

PF 16-MAR-2000; 2000WO-US06792.

PR 19-MAR-1999; 99US-0125362.

PR 10-DEC-1999; 99US-0169980.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Rosen GA, Ruben SM, Komatsoulis G;

DR WPI; 2000-579483/54.

DR N-PSDB; AAC74269.

PT Isolated nucleic acid molecule encoding a human secreted protein is
 PT used in preventing, treating or ameliorating a medical condition -

PS Disclosure; Page 69-70; 434pp; English.

CC The polynucleotide sequences given in AAC74223-C74279 encode the human
 CC secreted proteins represented in AAB39179-B39226. Sequences
 CC AAB39227-B39308 are alternative proteins encoded by the genes, and also
 CC protein sequences with which they share homology. The proteins have
 CC activities based on the tissues and cells in which they are expressed.
 CC Examples of activities include: immunosuppressive; antiarthritic;
 CC antirheumatic; antiproliferative; cytostatic; cardiant; vasotropic;
 CC cerebroprotective; neotropic; neuroprotective; antibacterial; virucide;
 CC fungicide; and ophthalmological. The human secreted proteins,
 CC polynucleotides, antagonists and agonists of the invention may be useful
 CC in the treatment, prevention, and/or diagnosis of various disease,
 CC disorders and conditions such as autoimmune diseases e.g. rheumatoid
 CC arthritis, hyperproliferative disorders e.g. neoplasms of the breast or
 CC liver, cardiovascular disorders e.g. cardiac arrest, cerebrovascular
 CC disorders e.g. cerebral ischemia, angiogenesis, nervous system disorders
 CC e.g. Alzheimer's disease, infections caused by bacteria, viruses and
 CC fungi and ocular disorders e.g. corneal infection. The polypeptides can
 CC also be used to aid wound healing and epithelial cell proliferation, to
 CC regenerate tissues, maintain organs before transplantation, in
 CC chemotaxis and as a food additive or preservative e.g. to increase
 CC storage capabilities. Sequences AAC74214-C74222 and AAB39178 are used
 CC during the isolation and characterisation of the genes of the invention.

Sequence 62 AA;

Query Match 0.7%; Score 7; DB 21; Length 62;
 Best Local Similarity 100.0%; Pred. No. 73;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 669 N1YLSFH 675
 11111111
 DB 28 n1ylsfch 34

RESULT 13

ID ABB44131 standard; Peptide; 63 AA.

AC ABB44131;

DT 04-FEB-2002 (first entry)

DE Peptide #11637 encoded by human foetal liver single exon probe.

KM Human; foetal liver; gene expression; single exon nucleic acid probe.

OS Homo sapiens.

PN WO200157277-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00669.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

PA (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI; 2001-483447/52.

PT Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human fetal liver -

PS Claim 27; SEQ ID NO 36766; 639pp + sequence listing; English.

CC The invention relates to a single exon nucleic acid probe for
 CC measuring human gene expression in a sample derived from human foetal
 CC liver. The single exon nucleic acid probes may be used for predicting,
 CC measuring and displaying gene expression in samples derived from human
 CC fetal liver. The present sequence is a peptide encoded by a single exon
 CC nucleic acid probe of the invention.

CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.

Sequence 63 AA;

Query Match 0.7%; Score 7; DB 22; Length 63;
 Best Local Similarity 100.0%; Pred. No. 75;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 SSRSSS 38
 11111111
 DB 27 ssrsgss 33

RESULT 14

ID ABB17718 standard; Protein; 63 AA.

XX ABB17718;
AC
XX
XX 23-JAN-2002 (first entry)
DE
XX
XX Human nervous system related polypeptide SEQ ID NO 6375.
XX
XX Human: nootropic; neuroprotective; cytosolic; dermatological; virocidic;
KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;
KW antiparkinsonian; antisticking; antianaemic; antiarthritic; cancer;
KW antileukemic; hepatotropic; cerebroprotective; antiinflammatory;
KW antiallergic; antidiabetic; antiulcer; anticonvulsant; antifungal;
KW antiparasitic; cardiac; immune disorder; cardiovascular disorder;
KW neurological disease; infection; nephrotropic; gene therapy; vaccine.
XX
OS Homo sapiens.
XX
XX W0200159063-A2.
PN
XX
XX 16-AUG-2001.
PD
XX
XX 17-JAN-2001: 2001WO-US01334.
PF
XX
XX 31-JAN-2000: 2000US-0179065.
PR 04-FEB-2000: 2000US-0180628.
PR 24-FEB-2000: 2000US-0184654.
PR 02-MAR-2000: 2000US-0186350.
PR 16-MAR-2000: 2000US-0189874.
PR 17-MAR-2000: 2000US-0190076.
PR 18-APR-2000: 2000US-0198123.
PR 19-MAY-2000: 2000US-0205515.
PR 07-JUN-2000: 2000US-0209467.
PR 28-JUN-2000: 2000US-0214886.
PR 30-JUN-2000: 2000US-0215135.
PR 07-JUL-2000: 2000US-0216647.
PR 07-JUL-2000: 2000US-0216880.
PR 11-JUL-2000: 2000US-0217487.
PR 11-JUL-2000: 2000US-0217496.
PR 14-JUL-2000: 2000US-0218290.
PR 26-JUL-2000: 2000US-0220963.
PR 26-JUL-2000: 2000US-0220964.
PR 14-AUG-2000: 2000US-0224518.
PR 14-AUG-2000: 2000US-0224519.
PR 14-AUG-2000: 2000US-0225213.
PR 14-AUG-2000: 2000US-0225214.
PR 14-AUG-2000: 2000US-0225266.
PR 14-AUG-2000: 2000US-0225267.
PR 14-AUG-2000: 2000US-0225268.
PR 14-AUG-2000: 2000US-0225270.
PR 14-AUG-2000: 2000US-0225447.
PR 14-AUG-2000: 2000US-0225757.
PR 14-AUG-2000: 2000US-0225758.
PR 14-AUG-2000: 2000US-0225759.
PR 18-AUG-2000: 2000US-0226279.
PR 22-AUG-2000: 2000US-0226681.
PR 22-AUG-2000: 2000US-0226868.
PR 22-AUG-2000: 2000US-0227182.
PR 23-AUG-2000: 2000US-0227009.
PR 30-AUG-2000: 2000US-0228924.
PR 01-SEP-2000: 2000US-0229287.
PR 01-SEP-2000: 2000US-0229343.
PR 01-SEP-2000: 2000US-0229344.
PR 01-SEP-2000: 2000US-0229345.
PR 05-SEP-2000: 2000US-0229509.
PR 05-SEP-2000: 2000US-0229513.
PR 06-SEP-2000: 2000US-0230437.
PR 06-SEP-2000: 2000US-0230438.
PR 08-SEP-2000: 2000US-0231242.
PR 08-SEP-2000: 2000US-0231243.
PR 08-SEP-2000: 2000US-0231244.
PR 08-SEP-2000: 2000US-0231413.
PR 08-SEP-2000: 2000US-0231414.
PR 08-SEP-2000: 2000US-0232080.

PR 08-SEP-2000: 2000US-0232081.
PR 12-SEP-2000: 2000US-0231968.
PR 14-SEP-2000: 2000US-0232397.
PR 14-SEP-2000: 2000US-0232398.
PR 14-SEP-2000: 2000US-0232399.
PR 14-SEP-2000: 2000US-0232400.
PR 14-SEP-2000: 2000US-0232401.
PR 14-SEP-2000: 2000US-0233063.
PR 14-SEP-2000: 2000US-0233064.
PR 14-SEP-2000: 2000US-0233065.
PR 21-SEP-2000: 2000US-0234223.
PR 21-SEP-2000: 2000US-0234274.
PR 25-SEP-2000: 2000US-0234997.
PR 25-SEP-2000: 2000US-0234998.
PR 26-SEP-2000: 2000US-0235484.
PR 27-SEP-2000: 2000US-0235834.
PR 27-SEP-2000: 2000US-0235836.
PR 29-SEP-2000: 2000US-0236367.
PR 29-SEP-2000: 2000US-0236367.
PR 29-SEP-2000: 2000US-0236368.
PR 29-SEP-2000: 2000US-0236369.
PR 29-SEP-2000: 2000US-0236370.
PR 02-OCT-2000: 2000US-0236802.
PR 02-OCT-2000: 2000US-0237037.
PR 02-OCT-2000: 2000US-0237038.
PR 02-OCT-2000: 2000US-0237040.
PR 13-OCT-2000: 2000US-0239935.
PR 13-OCT-2000: 2000US-0239937.
PR 20-OCT-2000: 2000US-0240960.
PR 20-OCT-2000: 2000US-0241785.
PR 20-OCT-2000: 2000US-0241786.
PR 20-OCT-2000: 2000US-0241787.
PR 20-OCT-2000: 2000US-0241808.
PR 20-OCT-2000: 2000US-0241809.
PR 20-OCT-2000: 2000US-0241826.
PR 01-NOV-2000: 2000US-0242221.
PR 08-NOV-2000: 2000US-0246474.
PR 08-NOV-2000: 2000US-0246475.
PR 08-NOV-2000: 2000US-0246476.
PR 08-NOV-2000: 2000US-0246477.
PR 08-NOV-2000: 2000US-0246478.
PR 08-NOV-2000: 2000US-0246523.
PR 08-NOV-2000: 2000US-0246524.
PR 08-NOV-2000: 2000US-0246525.
PR 08-NOV-2000: 2000US-0246526.
PR 08-NOV-2000: 2000US-0246527.
PR 08-NOV-2000: 2000US-0246528.
PR 08-NOV-2000: 2000US-0246532.
PR 08-NOV-2000: 2000US-0246609.
PR 08-NOV-2000: 2000US-0246610.
PR 08-NOV-2000: 2000US-0246611.
PR 08-NOV-2000: 2000US-0246613.
PR 17-NOV-2000: 2000US-0249207.
PR 17-NOV-2000: 2000US-0249208.
PR 17-NOV-2000: 2000US-0249209.
PR 17-NOV-2000: 2000US-0249210.
PR 17-NOV-2000: 2000US-0249211.
PR 17-NOV-2000: 2000US-0249212.
PR 17-NOV-2000: 2000US-0249213.
PR 17-NOV-2000: 2000US-0249214.
PR 17-NOV-2000: 2000US-0249215.
PR 17-NOV-2000: 2000US-0249216.
PR 17-NOV-2000: 2000US-0249217.
PR 17-NOV-2000: 2000US-0249218.
PR 17-NOV-2000: 2000US-0249244.
PR 17-NOV-2000: 2000US-0249245.
PR 17-NOV-2000: 2000US-0249264.
PR 17-NOV-2000: 2000US-0249265.
PR 17-NOV-2000: 2000US-0249297.
PR 17-NOV-2000: 2000US-0249299.
PR 17-NOV-2000: 2000US-0249300.

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 21, 2002, 14:51:15 ; Search time 17.88 seconds
(without alignments)
1467.175 Million cell updates/sec

Title: US-09-729-653-2_COPY_1_1074
Perfect score: 1074
Sequence: 1 HSLIGRCRGLGDGNNAVAC.....LYFEGEKRYLQAGKFFLLCG 1074

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 231628 seqs, 24425594 residues

Word size : 0

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep.*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep.*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep.*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep.*
5: /cgn2_6/ptodata/2/1aa/PCTUS.COMB.pep.*
6: /cgn2_6/ptodata/2/1aa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7	0.7	196	US-08-679-493A-185	Sequence 185, App
2	7	0.7	249	US-08-459-264-2	Sequence 2, Appl
3	7	0.7	249	US-08-459-264-3	Sequence 3, Appl
4	7	0.7	249	US-08-459-263-2	Sequence 2, Appl
5	7	0.7	249	US-08-459-263-3	Sequence 3, Appl
6	7	0.7	359	US-08-809-103B-2	Sequence 2, Appl
7	7	0.7	359	US-08-809-103B-4	Sequence 4, Appl
8	7	0.7	359	US-08-809-103B-6	Sequence 6, Appl
9	7	0.7	359	US-08-809-103B-8	Sequence 8, Appl
10	7	0.7	362	US-08-118-270-15	Sequence 15, Appl
11	7	0.7	362	PCT-US93-08528-15	Sequence 15, Appl
12	7	0.7	570	US-09-021-323-1	Sequence 1, Appl
13	7	0.7	574	US-08-907-166-12	Sequence 12, Appl
14	7	0.7	752	US-08-420-235B-21	Sequence 21, Appl
15	7	0.7	752	US-08-793-624-21	Sequence 21, Appl
16	7	0.7	752	PCT-US95-01194-21	Sequence 21, Appl
17	7	0.7	1612	US-08-169-927-2	Sequence 2, Appl
18	6	0.6	10	US-08-250-789A-67	Sequence 67, Appl
19	6	0.6	10	US-08-250-789A-76	Sequence 76, Appl
20	6	0.6	10	US-08-250-789A-83	Sequence 83, Appl
21	6	0.6	11	US-08-197-795-4	Sequence 4, Appl
22	6	0.6	11	PCT-US95-011721-4	Sequence 4, Appl
23	6	0.6	44	US-08-687-590-8	Sequence 8, Appl
24	6	0.6	44	US-08-687-590-9	Sequence 9, Appl
25	6	0.6	44	US-08-988-856B-6	Sequence 6, Appl
26	6	0.6	44	US-08-988-856B-16	Sequence 16, Appl
27	6	0.6	52	US-08-824-800D-12	Sequence 12, Appl

28	6	0.6	52	US-09-588-751-12	Sequence 12, Appl
29	6	0.6	56	US-08-867-087B-20	Sequence 20, Appl
30	6	0.6	64	US-09-627-376-6	Sequence 6, Appl
31	6	0.6	84	PCT-US93-08386-13	Sequence 13, Appl
32	6	0.6	94	US-08-466-368-5	Sequence 5, Appl
33	6	0.6	99	US-08-444-818-42	Sequence 42, Appl
34	6	0.6	100	US-08-672-850-13	Sequence 13, Appl
35	6	0.6	102	US-08-952-337-5	Sequence 5, Appl
36	6	0.6	102	US-08-894-173-75	Sequence 75, Appl
37	6	0.6	102	US-09-398-193-75	Sequence 75, Appl
38	6	0.6	103	US-08-472-171-2	Sequence 2, Appl
39	6	0.6	103	US-08-894-526-2	Sequence 2, Appl
40	6	0.6	103	US-09-013-047-2	Sequence 2, Appl
41	6	0.6	103	US-09-374-597-2	Sequence 2, Appl
42	6	0.6	103	US-08-857-076-35	Sequence 35, Appl
43	6	0.6	103	US-08-857-076-111	Sequence 111, App
44	6	0.6	104	US-09-012-515A-6	Sequence 6, Appl
45	6	0.6	104	US-08-360-144A-6	Sequence 6, Appl
46	6	0.6	104	PCT-US95-06722-6	Sequence 6, Appl
47	6	0.6	107	US-08-197-795-2	Sequence 2, Appl
48	6	0.6	107	US-08-336-618-13	Sequence 13, Appl
49	6	0.6	107	US-08-336-618-15	Sequence 15, Appl
50	6	0.6	107	US-08-963-601-1	Sequence 1, Appl
51	6	0.6	107	US-08-744-701-7	Sequence 7, Appl
52	6	0.6	107	US-08-894-173-67	Sequence 67, Appl
53	6	0.6	107	US-08-894-173-68	Sequence 68, Appl
54	6	0.6	107	US-08-894-173-80	Sequence 80, Appl
55	6	0.6	107	US-08-894-173-81	Sequence 81, Appl
56	6	0.6	107	US-09-398-193-67	Sequence 67, Appl
57	6	0.6	107	US-09-398-193-68	Sequence 68, Appl
58	6	0.6	107	US-09-398-193-80	Sequence 80, Appl
59	6	0.6	107	US-09-398-193-81	Sequence 81, Appl
60	6	0.6	107	PCT-US92-03993-2	Sequence 2, Appl
61	6	0.6	107	PCT-US92-03993-3	Sequence 3, Appl
62	6	0.6	107	PCT-US95-011721-2	Sequence 2, Appl
63	6	0.6	108	US-07-822-966B-5	Sequence 5, Appl
64	6	0.6	108	US-08-336-618-14	Sequence 14, Appl
65	6	0.6	108	US-08-803-899-3	Sequence 3, Appl
66	6	0.6	108	US-08-803-899-4	Sequence 4, Appl
67	6	0.6	108	US-08-961-083-166	Sequence 166, App
68	6	0.6	109	US-08-894-173-95	Sequence 95, Appl
69	6	0.6	109	US-09-398-193-95	Sequence 95, Appl
70	6	0.6	115	US-07-903-029-4	Sequence 4, Appl
71	6	0.6	115	US-07-903-029-5	Sequence 5, Appl
72	6	0.6	116	US-08-702-344-16	Sequence 16, Appl
73	6	0.6	117	US-08-444-818-44	Sequence 44, Appl
74	6	0.6	123	US-08-952-337-1	Sequence 1, Appl
75	6	0.6	124	US-08-449-045C-4	Sequence 4, Appl
76	6	0.6	124	US-08-435-605A-12	Sequence 12, Appl
77	6	0.6	124	US-09-012-097A-22	Sequence 22, Appl
78	6	0.6	128	US-08-687-816-20	Sequence 20, Appl
79	6	0.6	135	US-08-687-816-20	Sequence 20, Appl
80	6	0.6	135	US-09-138-614-20	Sequence 20, Appl
81	6	0.6	138	US-08-630-172-1	Sequence 1, Appl
82	6	0.6	138	US-09-375-419-1	Sequence 1, Appl
83	6	0.6	141	US-08-766-858A-46	Sequence 46, Appl
84	6	0.6	150	US-08-592-214A-14	Sequence 14, Appl
85	6	0.6	150	US-08-659-188-14	Sequence 14, Appl
86	6	0.6	150	US-08-655-427-14	Sequence 14, Appl
87	6	0.6	150	US-08-655-421-14	Sequence 14, Appl
88	6	0.6	150	US-09-149-976-14	Sequence 14, Appl
89	6	0.6	150	US-09-398-326-14	Sequence 14, Appl
90	6	0.6	152	US-08-672-850-6	Sequence 6, Appl
91	6	0.6	152	US-08-672-850-6	Sequence 6, Appl
92	6	0.6	152	US-08-672-850-12	Sequence 12, Appl
93	6	0.6	154	US-09-240-639-14	Sequence 14, Appl
94	6	0.6	159	US-08-844-086-4	Sequence 4, Appl
95	6	0.6	159	US-09-018-211-4	Sequence 4, Appl
96	6	0.6	160	US-08-961-539-4	Sequence 4, Appl
97	6	0.6	160	US-09-185-826-4	Sequence 4, Appl
98	6	0.6	165	US-09-335-409-16	Sequence 16, Appl
99	6	0.6	165	US-09-568-102-16	Sequence 16, Appl
100	6	0.6	185	US-09-567-969-16	Sequence 16, Appl

101	6	0.6	185	4	US-09-568-480-16	Sequence 16, Appl	174	6	0.6	295	3	US-08-459-504B-23	Sequence 23, Appl
102	6	0.6	185	4	US-09-568-486-16	Sequence 16, Appl	175	6	0.6	295	3	US-08-459-444-23	Sequence 23, Appl
103	6	0.6	185	4	US-09-568-472-16	Sequence 16, Appl	176	6	0.6	295	3	US-09-547-422-23	Sequence 23, Appl
104	6	0.6	187	4	US-09-194-185-1	Sequence 1, Appl	177	6	0.6	295	6	5233394-9	Patent No. 5233394
105	6	0.6	200	4	US-08-205-697A-9	Sequence 9, Appl	178	6	0.6	303	2	US-08-961-539-2	Sequence 2, Appl
106	6	0.6	200	4	US-08-702-525-9	Sequence 9, Appl	179	6	0.6	303	4	US-09-185-828-2	Sequence 2, Appl
107	6	0.6	200	4	PCT-US95-02576-9	Sequence 9, Appl	180	6	0.6	306	2	US-08-147-772-4	Sequence 4, Appl
108	6	0.6	203	2	US-08-284-391B-31	Sequence 31, Appl	181	6	0.6	306	2	US-08-456-104-8	Sequence 8, Appl
109	6	0.6	203	4	US-09-218-950-31	Sequence 31, Appl	182	6	0.6	306	2	US-08-101-624-25	Sequence 25, Appl
110	6	0.6	209	2	US-08-827-279-1	Sequence 1, Appl	183	6	0.6	306	3	US-08-153-262-4	Sequence 4, Appl
111	6	0.6	211	5	PCT-US94-04174-18	Sequence 18, Appl	184	6	0.6	306	3	US-08-479-744A-31	Sequence 31, Appl
112	6	0.6	214	4	US-08-205-697A-11	Sequence 11, Appl	185	6	0.6	306	4	US-08-280-757B-31	Sequence 31, Appl
113	6	0.6	214	4	US-08-702-525-11	Sequence 11, Appl	186	6	0.6	306	4	US-09-159-135-4	Sequence 4, Appl
114	6	0.6	214	5	PCT-US95-02576-11	Sequence 11, Appl	187	6	0.6	306	4	US-08-205-697A-17	Sequence 17, Appl
115	6	0.6	222	3	US-08-942-001-8	Sequence 8, Appl	188	6	0.6	306	4	US-08-702-525-17	Sequence 17, Appl
116	6	0.6	222	4	US-09-337-386-8	Sequence 8, Appl	189	6	0.6	306	4	US-09-450-798-4	Sequence 4, Appl
117	6	0.6	223	1	US-07-708-885B-3	Sequence 3, Appl	190	6	0.6	306	5	PCT-US95-02576-17	Sequence 17, Appl
118	6	0.6	223	1	US-07-714-386-3	Sequence 3, Appl	191	6	0.6	309	1	US-08-557-115-3	Sequence 3, Appl
119	6	0.6	223	1	US-07-708-888A-3	Sequence 3, Appl	192	6	0.6	309	5	PCT-US94-03869-3	Sequence 3, Appl
120	6	0.6	224	1	US-08-707-793A-6	Sequence 6, Appl	193	6	0.6	310	3	US-08-477-460B-6	Sequence 6, Appl
121	6	0.6	224	1	US-08-707-792A-6	Sequence 6, Appl	194	6	0.6	310	3	US-08-379-516-6	Sequence 6, Appl
122	6	0.6	227	2	US-08-440-517A-4	Sequence 4, Appl	195	6	0.6	310	4	US-09-329-916-6	Sequence 6, Appl
123	6	0.6	227	4	US-09-092-160-4	Sequence 4, Appl	196	6	0.6	310	4	US-08-485-372A-6	Sequence 6, Appl
124	6	0.6	236	4	US-09-095-758-12	Sequence 12, Appl	197	6	0.6	310	4	US-09-409-006A-6	Sequence 6, Appl
125	6	0.6	236	4	US-09-422-968-12	Sequence 12, Appl	198	6	0.6	310	5	PCT-US93-07422-6	Sequence 6, Appl
126	6	0.6	249	4	US-08-591-468-2	Sequence 2, Appl	199	6	0.6	318	1	US-08-247-809A-6	Sequence 6, Appl
127	6	0.6	249	5	PCT-US94-06430-2	Sequence 2, Appl	200	6	0.6	318	1	US-08-247-809A-6	Sequence 6, Appl
128	6	0.6	251	2	US-08-592-214A-12	Sequence 12, Appl	201	6	0.6	318	2	US-08-711-728-2	Sequence 2, Appl
129	6	0.6	251	3	US-08-659-188-12	Sequence 12, Appl	202	6	0.6	318	2	US-08-711-728-6	Sequence 6, Appl
130	6	0.6	251	3	US-08-655-227-12	Sequence 12, Appl	203	6	0.6	318	6	5233394-11	Patent No. 5233394
131	6	0.6	251	3	US-08-655-241-12	Sequence 12, Appl	204	6	0.6	320	4	US-08-205-697A-2	Sequence 2, Appl
132	6	0.6	251	3	US-09-149-976-12	Sequence 12, Appl	205	6	0.6	320	4	US-08-702-525-2	Sequence 2, Appl
133	6	0.6	251	4	US-09-398-326-12	Sequence 12, Appl	206	6	0.6	320	5	PCT-US95-02576-2	Sequence 2, Appl
134	6	0.6	258	4	US-08-495-484-13	Sequence 13, Appl	207	6	0.6	321	4	US-08-485-511A-4	Sequence 4, Appl
135	6	0.6	259	1	US-07-901-707-10	Sequence 10, Appl	208	6	0.6	322	2	US-08-701-191A-11	Sequence 11, Appl
136	6	0.6	259	1	US-07-988-430-10	Sequence 10, Appl	209	6	0.6	324	2	US-08-793-410-29	Sequence 29, Appl
137	6	0.6	259	1	US-08-425-336-10	Sequence 10, Appl	210	6	0.6	324	2	US-08-793-410-30	Sequence 30, Appl
138	6	0.6	259	1	US-08-488-113B-10	Sequence 10, Appl	211	6	0.6	325	3	US-09-015-754-2	Sequence 2, Appl
139	6	0.6	259	1	US-08-477-484B-10	Sequence 10, Appl	212	6	0.6	325	3	US-09-015-754-2	Sequence 2, Appl
140	6	0.6	259	2	US-08-646-360-10	Sequence 10, Appl	213	6	0.6	328	2	US-08-793-410-6	Sequence 6, Appl
141	6	0.6	259	4	US-08-839-765-10	Sequence 10, Appl	214	6	0.6	328	2	US-08-793-410-7	Sequence 7, Appl
142	6	0.6	259	4	US-09-136-389-10	Sequence 10, Appl	215	6	0.6	330	1	US-08-468-851-6	Sequence 6, Appl
143	6	0.6	259	5	PCT-US92-09487-10	Sequence 10, Appl	216	6	0.6	330	1	US-08-468-851-6	Sequence 6, Appl
144	6	0.6	260	1	US-08-378-761A-72	Sequence 72, Appl	217	6	0.6	330	1	US-08-310-357-6	Sequence 6, Appl
145	6	0.6	260	1	US-08-485-286-72	Sequence 72, Appl	218	6	0.6	330	1	US-08-468-851-6	Sequence 6, Appl
146	6	0.6	262	1	US-08-241-766-8	Sequence 8, Appl	219	6	0.6	330	2	US-08-468-851-6	Sequence 6, Appl
147	6	0.6	262	1	US-08-241-766-9	Sequence 9, Appl	220	6	0.6	335	2	US-08-379-556A-10	Sequence 10, Appl
148	6	0.6	266	1	US-08-247-809A-12	Sequence 12, Appl	221	6	0.6	336	1	US-07-755-009-3	Sequence 3, Appl
149	6	0.6	266	2	US-08-711-728-12	Sequence 12, Appl	222	6	0.6	337	4	US-09-044-404A-2	Sequence 2, Appl
150	6	0.6	267	2	US-08-667-809B-2	Sequence 2, Appl	223	6	0.6	338	4	US-09-723-153-4	Sequence 4, Appl
151	6	0.6	269	2	US-07-857-224B-38	Sequence 38, Appl	224	6	0.6	339	1	US-08-431-080-18	Sequence 18, Appl
152	6	0.6	278	2	US-08-865-203-6	Sequence 6, Appl	225	6	0.6	339	2	US-08-938-534-18	Sequence 18, Appl
153	6	0.6	278	2	US-07-849-420-6	Sequence 6, Appl	226	6	0.6	342	4	US-09-364-230-6	Sequence 6, Appl
154	6	0.6	278	4	US-09-253-854-6	Sequence 6, Appl	227	6	0.6	344	4	US-09-120-365-63	Sequence 63, Appl
155	6	0.6	278	4	US-08-955-424-6	Sequence 4, Appl	228	6	0.6	344	4	US-09-515-039-63	Sequence 63, Appl
156	6	0.6	279	2	US-09-135-658-4	Sequence 4, Appl	229	6	0.6	345	2	US-08-758-621-14	Sequence 14, Appl
157	6	0.6	279	4	US-09-330-970-5	Sequence 5, Appl	230	6	0.6	345	4	US-09-107-858-14	Sequence 14, Appl
158	6	0.6	279	6	5472855-6	Patent No. 5472855	231	6	0.6	346	4	US-09-352-990-23	Sequence 23, Appl
159	6	0.6	281	4	US-08-652-877-7	Sequence 7, Appl	232	6	0.6	349	4	US-09-006-353A-13	Sequence 13, Appl
160	6	0.6	281	4	US-08-476-515A-7	Sequence 7, Appl	233	6	0.6	349	4	US-09-161-241-8	Sequence 8, Appl
161	6	0.6	288	4	US-09-378-528-2	Sequence 2, Appl	234	6	0.6	350	4	US-09-161-241-9	Sequence 9, Appl
162	6	0.6	289	2	US-08-284-941-10	Sequence 10, Appl	235	6	0.6	355	2	US-08-846-762-93	Sequence 93, Appl
163	6	0.6	289	2	US-08-447-642-10	Sequence 10, Appl	236	6	0.6	359	1	US-07-662-005A-4	Sequence 4, Appl
164	6	0.6	289	4	US-09-236-503-10	Sequence 10, Appl	237	6	0.6	359	2	US-08-483-926A-8	Sequence 8, Appl
165	6	0.6	289	5	PCT-US93-02147A-10	Sequence 10, Appl	238	6	0.6	359	2	US-08-737-045-11	Sequence 11, Appl
166	6	0.6	291	3	US-08-358-171-2	Sequence 2, Appl	239	6	0.6	359	3	US-08-933-871B-1	Sequence 1, Appl
167	6	0.6	291	3	US-09-090-947-2	Sequence 2, Appl	240	6	0.6	359	3	US-09-476-919-1	Sequence 1, Appl
168	6	0.6	294	1	US-08-325-562-2	Sequence 2, Appl	241	6	0.6	359	4	US-09-347-798-12	Sequence 12, Appl
169	6	0.6	294	1	US-08-437-715-2	Sequence 2, Appl	242	6	0.6	359	4	US-08-780-311A-1	Sequence 1, Appl
170	6	0.6	295	1	US-07-951-715A-23	Sequence 23, Appl	243	6	0.6	361	4	US-09-108-020-47	Sequence 47, Appl
171	6	0.6	295	2	US-08-459-448A-23	Sequence 23, Appl	244	6	0.6	364	2	US-08-483-926A-10	Sequence 10, Appl
172	6	0.6	295	3	US-08-459-595A-23	Sequence 23, Appl	245	6	0.6	364	2	US-08-737-045-10	Sequence 10, Appl
173	6	0.6	295	3	US-09-002-298-5	Sequence 5, Appl	246	6	0.6	367	2	US-08-990-379-6	Sequence 6, Appl

247	6	0.6	375	1	US-08-464-523B-24	Sequence 24, Appl	320	6	0.6	428	3	US-08-883-610A-2	Sequence 2, Appl
248	6	0.6	376	2	US-08-558-269-10	Sequence 10, Appl	321	6	0.6	428	4	US-08-936-094A-2	Sequence 2, Appl
249	6	0.6	376	4	US-09-410-882-10	Sequence 10, Appl	322	6	0.6	428	4	US-09-347-833-6	Sequence 6, Appl
250	6	0.6	377	1	US-08-188-277B-2	Sequence 2, Appl	323	6	0.6	429	2	US-08-310-912A-109	Sequence 109, App
251	6	0.6	377	1	US-08-188-277B-4	Sequence 4, Appl	324	6	0.6	429	4	US-09-301-0485-109	Sequence 109, App
252	6	0.6	377	2	US-08-429-964-78	Sequence 78, Appl	325	6	0.6	429	5	PCT-US95-08589-109	Sequence 109, App
253	6	0.6	377	2	US-08-429-964-80	Sequence 80, Appl	326	6	0.6	432	3	US-08-477-460B-1	Sequence 2, Appl
254	6	0.6	378	1	US-08-707-793A-4	Sequence 4, Appl	327	6	0.6	432	3	US-08-379-516-2	Sequence 2, Appl
255	6	0.6	378	1	US-08-707-792A-4	Sequence 4, Appl	328	6	0.6	432	3	US-09-329-916-2	Sequence 2, Appl
256	6	0.6	380	2	US-08-846-762A-86	Sequence 86, Appl	329	6	0.6	432	4	US-08-485-372A-2	Sequence 2, Appl
257	6	0.6	383	1	US-08-464-523B-31	Sequence 31, Appl	330	6	0.6	432	4	US-09-409-006A-2	Sequence 2, Appl
258	6	0.6	383	2	US-08-558-269-6	Sequence 6, Appl	331	6	0.6	432	5	PCT-US93-07422-2	Sequence 2, Appl
259	6	0.6	383	4	US-09-410-882-6	Sequence 6, Appl	332	6	0.6	433	2	US-08-867-149-1	Sequence 1, Appl
260	6	0.6	384	1	US-08-707-793A-5	Sequence 5, Appl	333	6	0.6	433	2	US-08-808-374-1	Sequence 1, Appl
261	6	0.6	384	1	US-08-707-792A-5	Sequence 5, Appl	334	6	0.6	433	3	US-09-100-409A-1	Sequence 1, Appl
262	6	0.6	386	2	US-08-858-052-4	Sequence 4, Appl	335	6	0.6	433	4	US-09-400-208B-5	Sequence 5, Appl
263	6	0.6	386	3	US-09-200-284-4	Sequence 4, Appl	336	6	0.6	433	6	517183B-13	Patent No. 517183B
264	6	0.6	386	4	US-08-965-762-5	Sequence 5, Appl	337	6	0.6	434	1	US-08-336-311-4	Sequence 4, Appl
265	6	0.6	388	4	US-09-188-930-775	Sequence 275, App	338	6	0.6	434	1	US-08-162-392-4	Sequence 4, Appl
266	6	0.6	394	1	US-07-914-281-4	Sequence 4, Appl	339	6	0.6	434	1	US-08-643-300-4	Sequence 4, Appl
267	6	0.6	394	1	US-08-393-246-4	Sequence 4, Appl	340	6	0.6	434	1	US-08-643-297-4	Sequence 4, Appl
268	6	0.6	394	1	US-08-525-058A-4	Sequence 4, Appl	341	6	0.6	434	1	US-08-643-298-4	Sequence 4, Appl
269	6	0.6	394	2	US-08-636-731-4	Sequence 4, Appl	342	6	0.6	434	2	US-08-643-301-4	Sequence 4, Appl
270	6	0.6	394	3	US-08-466-368-2	Sequence 2, Appl	343	6	0.6	434	3	US-08-457-918-4	Sequence 4, Appl
271	6	0.6	394	4	US-09-042-531-4	Sequence 4, Appl	344	6	0.6	434	5	PCT-US94-14095-4	Sequence 4, Appl
272	6	0.6	394	5	PCT-US91-00899-9	Sequence 9, Appl	345	6	0.6	439	3	US-08-993-359-24	Sequence 24, Appl
273	6	0.6	394	6	5223418-2	Patent No. 5223418	346	6	0.6	439	3	US-09-321-654-2	Sequence 2, Appl
274	6	0.6	398	2	US-08-284-391B-29	Sequence 29, Appl	347	6	0.6	439	3	US-08-989-358A-2	Sequence 2, Appl
275	6	0.6	398	4	US-09-218-950-29	Sequence 29, Appl	348	6	0.6	439	4	US-09-457-046B-68	Sequence 68, Appl
276	6	0.6	399	4	US-08-765-907A-10	Sequence 10, Appl	349	6	0.6	441	4	US-09-171-461-26	Sequence 26, Appl
277	6	0.6	402	1	US-08-236-311-1	Sequence 1, Appl	350	6	0.6	443	3	US-09-120-365-69	Sequence 69, Appl
278	6	0.6	402	1	US-08-553-703A-1	Sequence 1, Appl	351	6	0.6	443	4	US-09-515-039-69	Sequence 69, Appl
279	6	0.6	402	1	US-08-553-703A-4	Sequence 4, Appl	352	6	0.6	448	3	US-09-039-555B-15	Sequence 15, Appl
280	6	0.6	402	2	US-09-006-021-1	Sequence 1, Appl	353	6	0.6	458	3	US-08-466-368-4	Sequence 4, Appl
281	6	0.6	402	2	US-09-006-021-4	Sequence 4, Appl	354	6	0.6	458	6	522394-7	Patent No. 522394
282	6	0.6	402	3	US-08-457-918-1	Sequence 1, Appl	355	6	0.6	461	1	US-07-796-361A-13	Sequence 13, Appl
283	6	0.6	404	1	US-08-203-716-2	Sequence 2, Appl	356	6	0.6	462	1	US-07-612-673-4	Sequence 4, Appl
284	6	0.6	404	1	US-08-242-663A-2	Sequence 2, Appl	357	6	0.6	462	1	US-08-539-666-4	Sequence 4, Appl
285	6	0.6	404	1	US-08-440-179-2	Sequence 2, Appl	358	6	0.6	462	2	US-08-417-495-5	Sequence 5, Appl
286	6	0.6	404	2	US-08-450-130A-1	Sequence 1, Appl	359	6	0.6	462	2	US-08-284-391B-5	Sequence 5, Appl
287	6	0.6	404	2	US-08-391-916A-2	Sequence 2, Appl	360	6	0.6	462	4	US-09-218-950-5	Sequence 5, Appl
288	6	0.6	404	2	US-08-573-890-2	Sequence 2, Appl	361	6	0.6	462	5	PCT-US92-01785-5	Sequence 5, Appl
289	6	0.6	404	2	US-08-394-189B-4	Sequence 4, Appl	362	6	0.6	462	5	PCT-US95-00454-5	Sequence 5, Appl
290	6	0.6	404	3	US-08-450-362A-1	Sequence 1, Appl	363	6	0.6	464	2	US-08-878-989-4	Sequence 4, Appl
291	6	0.6	404	3	US-08-258-287B-39	Sequence 39, Appl	364	6	0.6	464	3	US-09-056-783-2	Sequence 2, Appl
292	6	0.6	404	3	US-08-368-704C-39	Sequence 39, Appl	365	6	0.6	464	4	US-09-272-796-4	Sequence 4, Appl
293	6	0.6	404	3	US-08-954-536-18	Sequence 18, Appl	366	6	0.6	464	4	US-08-487-550-4	Sequence 4, Appl
294	6	0.6	404	4	US-09-039-657-2	Sequence 2, Appl	367	6	0.6	477	1	US-07-735-065-2	Sequence 2, Appl
295	6	0.6	404	4	US-08-748-547-2	Sequence 2, Appl	368	6	0.6	477	2	US-08-469-202-12	Sequence 12, Appl
296	6	0.6	404	4	US-08-908-436-3	Sequence 3, Appl	369	6	0.6	477	2	US-08-484-434C-12	Sequence 12, Appl
297	6	0.6	404	4	US-09-248-179-2	Sequence 2, Appl	370	6	0.6	480	2	US-08-570-227A-2	Sequence 2, Appl
298	6	0.6	404	4	US-09-069-023-30	Sequence 30, Appl	371	6	0.6	480	4	US-09-077-891-2	Sequence 2, Appl
299	6	0.6	404	5	PCT-US93-05705-4	Sequence 4, Appl	372	6	0.6	484	3	US-08-369-822C-21	Sequence 21, Appl
300	6	0.6	404	5	PCT-US95-06132-2	Sequence 2, Appl	373	6	0.6	484	3	US-08-582-776C-36	Sequence 36, Appl
301	6	0.6	404	5	PCT-US95-07619-1	Sequence 1, Appl	374	6	0.6	484	3	US-08-434-831B-33	Sequence 33, Appl
302	6	0.6	404	5	PCT-US96-10521-13	Sequence 13, Appl	375	6	0.6	491	2	US-08-687-916-23	Sequence 23, Appl
303	6	0.6	406	5	PCT-US94-02107-2	Sequence 2, Appl	376	6	0.6	491	2	US-09-001-826-23	Sequence 23, Appl
304	6	0.6	410	3	US-08-630-172-17	Sequence 17, Appl	377	6	0.6	491	4	US-09-138-614-23	Sequence 23, Appl
305	6	0.6	410	4	US-09-375-419-17	Sequence 17, Appl	378	6	0.6	492	2	US-08-644-271-32	Sequence 32, Appl
306	6	0.6	413	1	US-08-444-734A-5	Sequence 5, Appl	379	6	0.6	494	4	US-09-019-095A-24	Sequence 24, Appl
307	6	0.6	413	1	US-08-087-772A-17	Sequence 17, Appl	380	6	0.6	495	1	US-08-532-065B-2	Sequence 2, Appl
308	6	0.6	413	2	US-08-467-568-12	Sequence 12, Appl	381	6	0.6	495	2	US-08-687-916-22	Sequence 22, Appl
309	6	0.6	413	3	US-09-030-582-12	Sequence 12, Appl	382	6	0.6	495	4	US-09-138-614-22	Sequence 22, Appl
310	6	0.6	413	5	PCT-US91-00909-4	Sequence 4, Appl	383	6	0.6	496	2	US-08-949-637-2	Sequence 2, Appl
311	6	0.6	416	1	US-08-464-523B-33	Sequence 33, Appl	384	6	0.6	496	4	US-09-231-488-2	Sequence 2, Appl
312	6	0.6	417	1	US-08-464-523B-23	Sequence 23, Appl	385	6	0.6	499	4	US-08-988-856B-2	Sequence 2, Appl
313	6	0.6	418	4	US-09-198-603C-8	Sequence 8, Appl	386	6	0.6	499	4	US-08-988-856B-4	Sequence 4, Appl
314	6	0.6	423	1	US-08-464-523B-25	Sequence 25, Appl	387	6	0.6	499	5	PCT-US95-01806-2	Sequence 2, Appl
315	6	0.6	424	4	US-09-627-376-8	Sequence 8, Appl	388	6	0.6	499	5	PCT-US95-01806-4	Sequence 4, Appl
316	6	0.6	425	4	US-09-230-371A-30	Sequence 30, Appl	389	6	0.6	504	1	US-07-853-985A-6	Sequence 6, Appl
317	6	0.6	426	4	US-08-961-083-48	Sequence 48, Appl	390	6	0.6	504	1	US-07-681-703B-6	Sequence 6, Appl
318	6	0.6	427	3	US-09-196-857-2	Sequence 2, Appl	391	6	0.6	504	1	US-08-184-236-6	Sequence 6, Appl
319	6	0.6	428	2	US-08-717-587A-2	Sequence 2, Appl	392	6	0.6	504	2	US-08-407-410B-6	Sequence 6, Appl

393	6	0.6	504	2	US-08-485-500-6	Sequence 6, Appl1	466	6	0.6	621	3	US-08-604-789B-4	Sequence 4, Appl1
394	6	0.6	504	5	PCT-US91-02370-6	Sequence 6, Appl1	467	6	0.6	621	4	US-09-312-721A-4	Sequence 4, Appl1
395	6	0.6	504	5	PCT-US94-04174-6	Sequence 6, Appl1	468	6	0.6	626	2	US-08-596-300A-7	Sequence 7, Appl1
396	6	0.6	523	1	US-08-100-247-2	Sequence 2, Appl1	469	6	0.6	626	2	US-08-596-300A-14	Sequence 14, Appl1
397	6	0.6	523	1	US-08-483-146A-2	Sequence 2, Appl1	470	6	0.6	651	1	US-08-315-468-6	Sequence 6, Appl1
398	6	0.6	523	1	US-08-232-513A-3	Sequence 3, Appl1	471	6	0.6	651	3	US-08-996-441B-52	Sequence 52, Appl1
399	6	0.6	523	1	US-08-484-594A-2	Sequence 2, Appl1	472	6	0.6	651	3	US-08-996-441B-56	Sequence 56, Appl1
400	6	0.6	530	1	US-08-307-499-29	Sequence 29, Appl1	473	6	0.6	651	3	US-08-996-441B-58	Sequence 58, Appl1
401	6	0.6	530	3	US-08-477-460B-4	Sequence 4, Appl1	474	6	0.6	651	3	US-08-993-722A-52	Sequence 52, Appl1
402	6	0.6	530	3	US-08-379-516-4	Sequence 4, Appl1	475	6	0.6	651	3	US-08-993-722A-56	Sequence 56, Appl1
403	6	0.6	530	4	US-09-329-916-4	Sequence 4, Appl1	476	6	0.6	651	3	US-08-993-722A-58	Sequence 58, Appl1
404	6	0.6	530	4	US-08-485-372A-4	Sequence 4, Appl1	477	6	0.6	651	3	US-08-993-170A-52	Sequence 52, Appl1
405	6	0.6	530	4	US-09-289-268-29	Sequence 29, Appl1	478	6	0.6	651	3	US-08-993-170A-56	Sequence 56, Appl1
406	6	0.6	530	4	US-09-409-006A-4	Sequence 4, Appl1	479	6	0.6	651	3	US-08-993-170A-58	Sequence 58, Appl1
407	6	0.6	530	5	PCT-US93-07422-4	Sequence 4, Appl1	480	6	0.6	651	3	US-08-993-775B-52	Sequence 52, Appl1
408	6	0.6	532	2	US-08-417-495-6	Sequence 6, Appl1	481	6	0.6	651	3	US-08-993-775B-56	Sequence 56, Appl1
409	6	0.6	532	2	US-08-284-391B-6	Sequence 6, Appl1	482	6	0.6	651	3	US-08-993-775B-58	Sequence 58, Appl1
410	6	0.6	532	2	US-08-560-916-8	Sequence 8, Appl1	483	6	0.6	651	4	US-07-941-650A-4	Sequence 4, Appl1
411	6	0.6	532	2	US-08-676-841-8	Sequence 8, Appl1	484	6	0.6	652	3	US-08-996-441B-2	Sequence 2, Appl1
412	6	0.6	532	4	US-09-218-950-6	Sequence 6, Appl1	485	6	0.6	652	3	US-08-996-441B-4	Sequence 4, Appl1
413	6	0.6	532	5	PCT-US92-01785-6	Sequence 6, Appl1	486	6	0.6	652	3	US-08-996-441B-6	Sequence 6, Appl1
414	6	0.6	532	5	PCT-US95-00454-6	Sequence 6, Appl1	487	6	0.6	652	3	US-08-996-441B-8	Sequence 8, Appl1
415	6	0.6	535	2	US-08-933-750C-20	Sequence 20, Appl1	488	6	0.6	652	3	US-08-996-441B-10	Sequence 10, Appl1
416	6	0.6	535	4	US-09-234-613-20	Sequence 20, Appl1	489	6	0.6	652	3	US-08-996-441B-12	Sequence 12, Appl1
417	6	0.6	541	4	US-08-867-611-8	Sequence 8, Appl1	490	6	0.6	652	3	US-08-996-441B-14	Sequence 14, Appl1
418	6	0.6	541	4	US-08-687-590-28	Sequence 28, Appl1	491	6	0.6	652	3	US-08-996-441B-16	Sequence 16, Appl1
419	6	0.6	541	4	US-09-311-311C-25	Sequence 25, Appl1	492	6	0.6	652	3	US-08-996-441B-18	Sequence 18, Appl1
420	6	0.6	541	5	PCT-US92-06965A-13	Sequence 13, Appl1	493	6	0.6	652	3	US-08-996-441B-20	Sequence 20, Appl1
421	6	0.6	543	2	US-08-469-412A-7	Sequence 7, Appl1	494	6	0.6	652	3	US-08-996-441B-22	Sequence 22, Appl1
422	6	0.6	543	4	US-09-021-715-7	Sequence 7, Appl1	495	6	0.6	652	3	US-08-996-441B-24	Sequence 24, Appl1
423	6	0.6	544	1	US-08-387-156-10	Sequence 10, Appl1	496	6	0.6	652	3	US-08-996-441B-26	Sequence 26, Appl1
424	6	0.6	544	2	US-08-694-865-10	Sequence 10, Appl1	497	6	0.6	652	3	US-08-996-441B-28	Sequence 28, Appl1
425	6	0.6	544	2	US-08-878-748-10	Sequence 10, Appl1	498	6	0.6	652	3	US-08-996-441B-30	Sequence 30, Appl1
426	6	0.6	544	2	US-08-587-680A-25	Sequence 25, Appl1	499	6	0.6	652	3	US-08-996-441B-32	Sequence 32, Appl1
427	6	0.6	544	3	US-09-124-491-10	Sequence 10, Appl1	500	6	0.6	652	3	US-08-996-441B-34	Sequence 34, Appl1
428	6	0.6	544	4	US-08-687-590-30	Sequence 30, Appl1	501	6	0.6	652	3	US-08-996-441B-36	Sequence 36, Appl1
429	6	0.6	545	4	US-08-990-114-1	Sequence 1, Appl1	502	6	0.6	652	3	US-08-996-441B-38	Sequence 38, Appl1
430	6	0.6	545	4	US-09-241-333-1	Sequence 1, Appl1	503	6	0.6	652	3	US-08-996-441B-40	Sequence 40, Appl1
431	6	0.6	548	2	US-08-469-412A-2	Sequence 2, Appl1	504	6	0.6	652	3	US-08-996-441B-42	Sequence 42, Appl1
432	6	0.6	548	4	US-09-021-715-2	Sequence 2, Appl1	505	6	0.6	652	3	US-08-996-441B-44	Sequence 44, Appl1
433	6	0.6	556	4	US-08-687-590-24	Sequence 24, Appl1	506	6	0.6	652	3	US-08-996-441B-46	Sequence 46, Appl1
434	6	0.6	559	1	US-08-424-788-3	Sequence 3, Appl1	507	6	0.6	652	3	US-08-996-441B-48	Sequence 48, Appl1
435	6	0.6	557	1	US-08-374-155A-16	Sequence 16, Appl1	508	6	0.6	652	3	US-08-996-441B-50	Sequence 50, Appl1
436	6	0.6	557	1	US-08-785-396-16	Sequence 16, Appl1	509	6	0.6	652	3	US-08-996-441B-54	Sequence 54, Appl1
437	6	0.6	575	1	US-08-403-866-7	Sequence 7, Appl1	510	6	0.6	652	3	US-08-996-441B-60	Sequence 60, Appl1
438	6	0.6	575	1	US-08-424-788-2	Sequence 2, Appl1	511	6	0.6	652	3	US-08-996-441B-62	Sequence 62, Appl1
439	6	0.6	575	1	US-08-110-683-4	Sequence 4, Appl1	512	6	0.6	652	3	US-08-996-441B-64	Sequence 64, Appl1
440	6	0.6	575	2	US-08-417-495-4	Sequence 4, Appl1	513	6	0.6	652	3	US-08-996-441B-66	Sequence 66, Appl1
441	6	0.6	575	2	US-08-284-391B-4	Sequence 4, Appl1	514	6	0.6	652	3	US-08-996-441B-68	Sequence 68, Appl1
442	6	0.6	575	2	US-08-477-166-4	Sequence 4, Appl1	515	6	0.6	652	3	US-08-996-441B-68	Sequence 68, Appl1
443	6	0.6	575	2	US-08-472-097-4	Sequence 4, Appl1	516	6	0.6	652	3	US-08-996-441B-108	Sequence 108, App
444	6	0.6	575	5	PCT-US92-01785-4	Sequence 4, Appl1	517	6	0.6	652	3	US-08-996-441B-110	Sequence 110, App
445	6	0.6	575	5	PCT-US93-11638-4	Sequence 4, Appl1	518	6	0.6	652	3	US-08-996-441B-111	Sequence 111, App
446	6	0.6	575	5	PCT-US95-00454-4	Sequence 4, Appl1	519	6	0.6	652	3	US-08-993-722A-2	Sequence 2, Appl1
447	6	0.6	591	2	US-08-736-770-5	Sequence 5, Appl1	520	6	0.6	652	3	US-08-993-722A-4	Sequence 4, Appl1
448	6	0.6	593	2	US-08-591-079-8	Sequence 8, Appl1	521	6	0.6	652	3	US-08-993-722A-6	Sequence 6, Appl1
449	6	0.6	593	2	US-08-591-079-10	Sequence 10, Appl1	522	6	0.6	652	3	US-08-993-722A-8	Sequence 8, Appl1
450	6	0.6	595	3	US-08-604-789B-3	Sequence 3, Appl1	523	6	0.6	652	3	US-08-993-722A-10	Sequence 10, Appl1
451	6	0.6	595	3	US-09-312-721A-3	Sequence 3, Appl1	524	6	0.6	652	3	US-08-993-722A-12	Sequence 12, Appl1
452	6	0.6	600	5	PCT-US95-10166-2	Sequence 2, Appl1	525	6	0.6	652	3	US-08-993-722A-14	Sequence 14, Appl1
453	6	0.6	601	1	US-08-446-038B-16	Sequence 16, Appl1	526	6	0.6	652	3	US-08-993-722A-16	Sequence 16, Appl1
454	6	0.6	601	1	US-08-446-010B-16	Sequence 16, Appl1	527	6	0.6	652	3	US-08-993-722A-18	Sequence 18, Appl1
455	6	0.6	601	1	US-08-805-445-16	Sequence 16, Appl1	528	6	0.6	652	3	US-08-993-722A-20	Sequence 20, Appl1
456	6	0.6	601	2	US-08-064-067D-16	Sequence 16, Appl1	529	6	0.6	652	3	US-08-993-722A-22	Sequence 22, Appl1
457	6	0.6	601	2	US-09-066-208-16	Sequence 16, Appl1	530	6	0.6	652	3	US-08-993-722A-24	Sequence 24, Appl1
458	6	0.6	603	3	US-08-482-677-8	Sequence 8, Appl1	531	6	0.6	652	3	US-08-993-722A-26	Sequence 26, Appl1
459	6	0.6	604	2	US-08-635-137-2	Sequence 2, Appl1	532	6	0.6	652	3	US-08-993-722A-30	Sequence 30, Appl1
460	6	0.6	604	4	US-09-136-981-2	Sequence 2, Appl1	533	6	0.6	652	3	US-08-993-722A-32	Sequence 32, Appl1
461	6	0.6	615	4	US-09-462-844-3	Sequence 3, Appl1	534	6	0.6	652	3	US-08-993-722A-34	Sequence 34, Appl1
463	6	0.6	620	4	US-09-232-200-61	Sequence 61, Appl1	535	6	0.6	652	3	US-08-993-722A-36	Sequence 36, Appl1
464	6	0.6	620	4	US-09-232-197-61	Sequence 61, Appl1	536	6	0.6	652	3	US-08-993-722A-38	Sequence 38, Appl1
465	6	0.6	620	4	US-09-232-201-61	Sequence 61, Appl1	538	6	0.6	652	3	US-08-993-722A-40	Sequence 40, Appl1

539	6	0.6	652	3	US-08-993-722A-42	Sequence 42, App1	612	6	0.6	652	3	US-08-993-775B-48	Sequence 48, App1
540	6	0.6	652	3	US-08-993-722A-44	Sequence 44, App1	613	6	0.6	652	3	US-08-993-775B-50	Sequence 50, App1
541	6	0.6	652	3	US-08-993-722A-46	Sequence 46, App1	614	6	0.6	652	3	US-08-993-775B-54	Sequence 54, App1
542	6	0.6	652	3	US-08-993-722A-48	Sequence 48, App1	615	6	0.6	652	3	US-08-993-775B-60	Sequence 60, App1
543	6	0.6	652	3	US-08-993-722A-50	Sequence 50, App1	616	6	0.6	652	3	US-08-993-775B-62	Sequence 62, App1
544	6	0.6	652	3	US-08-993-722A-54	Sequence 54, App1	617	6	0.6	652	3	US-08-993-775B-64	Sequence 64, App1
545	6	0.6	652	3	US-08-993-722A-60	Sequence 60, App1	618	6	0.6	652	3	US-08-993-775B-66	Sequence 66, App1
546	6	0.6	652	3	US-08-993-722A-62	Sequence 62, App1	619	6	0.6	652	3	US-08-993-775B-68	Sequence 68, App1
547	6	0.6	652	3	US-08-993-722A-64	Sequence 64, App1	620	6	0.6	652	3	US-08-993-775B-98	Sequence 98, App1
548	6	0.6	652	3	US-08-993-722A-66	Sequence 66, App1	621	6	0.6	652	3	US-08-993-775B-108	Sequence 108, App
549	6	0.6	652	3	US-08-993-722A-68	Sequence 68, App1	622	6	0.6	652	3	US-08-993-775B-110	Sequence 110, App
550	6	0.6	652	3	US-08-993-722A-98	Sequence 98, App1	623	6	0.6	652	3	US-08-993-775B-111	Sequence 111, App
551	6	0.6	652	3	US-08-993-722A-108	Sequence 108, App	624	6	0.6	652	5	PCT-US92-00040-2	Patent No. 5187091
552	6	0.6	652	3	US-08-993-722A-110	Sequence 110, App	625	6	0.6	652	6	US-08-996-441B-100	Sequence 100, App
553	6	0.6	652	3	US-08-993-722A-111	Sequence 111, App	626	6	0.6	653	3	US-08-993-722A-100	Sequence 100, App
554	6	0.6	652	3	US-08-993-170A-2	Sequence 2, App11	627	6	0.6	653	3	US-08-993-722A-100	Sequence 100, App
555	6	0.6	652	3	US-08-993-170A-4	Sequence 4, App11	628	6	0.6	653	3	US-08-993-170A-100	Sequence 100, App
556	6	0.6	652	3	US-08-993-170A-6	Sequence 6, App11	629	6	0.6	653	3	US-08-993-775B-100	Sequence 100, App
557	6	0.6	652	3	US-08-993-170A-8	Sequence 8, App11	630	6	0.6	655	4	US-09-347-833-4	Sequence 4, App11
558	6	0.6	652	3	US-08-993-170A-10	Sequence 10, App1	631	6	0.6	659	1	US-08-136-277-3	Sequence 3, App11
559	6	0.6	652	3	US-08-993-170A-12	Sequence 12, App1	632	6	0.6	659	2	US-08-479-403-3	Sequence 3, App11
560	6	0.6	652	3	US-08-993-170A-14	Sequence 14, App1	633	6	0.6	659	3	US-08-835-734-3	Sequence 3, App11
561	6	0.6	652	3	US-08-993-170A-16	Sequence 16, App1	634	6	0.6	659	3	US-08-996-441B-112	Sequence 112, App
562	6	0.6	652	3	US-08-993-170A-18	Sequence 18, App1	635	6	0.6	659	3	US-08-993-722A-112	Sequence 112, App
563	6	0.6	652	3	US-08-993-170A-20	Sequence 20, App1	636	6	0.6	659	3	US-08-993-170A-112	Sequence 112, App
564	6	0.6	652	3	US-08-993-170A-22	Sequence 22, App1	637	6	0.6	659	3	US-08-993-775B-112	Sequence 112, App
565	6	0.6	652	3	US-08-993-170A-24	Sequence 24, App1	638	6	0.6	675	4	US-09-171-878-1	Sequence 1, App11
566	6	0.6	652	3	US-08-993-170A-26	Sequence 26, App1	639	6	0.6	676	4	US-09-061-768A-2	Sequence 2, App11
567	6	0.6	652	3	US-08-993-170A-28	Sequence 28, App1	640	6	0.6	682	1	US-07-998-003A-107	Sequence 107, App
568	6	0.6	652	3	US-08-993-170A-30	Sequence 30, App1	641	6	0.6	682	1	US-08-453-748B-107	Sequence 107, App
569	6	0.6	652	3	US-08-993-170A-32	Sequence 32, App1	642	6	0.6	682	1	US-08-453-695A-107	Sequence 107, App
570	6	0.6	652	3	US-08-993-170A-34	Sequence 34, App1	643	6	0.6	682	1	US-08-268-161A-107	Sequence 107, App
571	6	0.6	652	3	US-08-993-170A-36	Sequence 36, App1	644	6	0.6	682	2	US-08-453-702A-107	Sequence 107, App
572	6	0.6	652	3	US-08-993-170A-38	Sequence 38, App1	645	6	0.6	682	4	US-09-099-639A-107	Sequence 107, App
573	6	0.6	652	3	US-08-993-170A-40	Sequence 40, App1	646	6	0.6	682	5	PCT-US93-12588-107	Sequence 107, App
574	6	0.6	652	3	US-08-993-170A-42	Sequence 42, App1	647	6	0.6	682	5	PCT-US93-08071-107	Sequence 107, App
575	6	0.6	652	3	US-08-993-170A-44	Sequence 44, App1	648	6	0.6	685	3	US-08-947-965-72	Sequence 72, App1
576	6	0.6	652	3	US-08-993-170A-46	Sequence 46, App1	649	6	0.6	686	3	US-08-947-965-73	Sequence 73, App1
577	6	0.6	652	3	US-08-993-170A-48	Sequence 48, App1	650	6	0.6	694	3	US-08-559-397A-31	Sequence 31, App1
578	6	0.6	652	3	US-08-993-170A-50	Sequence 50, App1	651	6	0.6	699	2	US-08-694-865-16	Sequence 16, App1
579	6	0.6	652	3	US-08-993-170A-54	Sequence 54, App1	652	6	0.6	699	2	US-09-124-491-16	Sequence 16, App1
580	6	0.6	652	3	US-08-993-170A-60	Sequence 60, App1	653	6	0.6	718	1	US-08-190-802A-44	Sequence 44, App1
581	6	0.6	652	3	US-08-993-170A-62	Sequence 62, App1	654	6	0.6	718	4	US-08-473-089A-44	Sequence 44, App1
582	6	0.6	652	3	US-08-993-170A-64	Sequence 64, App1	655	6	0.6	718	4	US-08-473-089A-44	Sequence 44, App1
583	6	0.6	652	3	US-08-993-170A-66	Sequence 66, App1	656	6	0.6	718	4	US-08-989-385A-19	Sequence 19, App1
584	6	0.6	652	3	US-08-993-170A-68	Sequence 68, App1	657	6	0.6	718	4	US-08-989-385A-19	Sequence 19, App1
585	6	0.6	652	3	US-08-993-170A-98	Sequence 98, App1	658	6	0.6	721	4	US-09-036-987A-24	Sequence 24, App1
586	6	0.6	652	3	US-08-993-170A-108	Sequence 108, App	659	6	0.6	751	4	US-09-370-740-24	Sequence 24, App1
587	6	0.6	652	3	US-08-993-170A-110	Sequence 110, App	660	6	0.6	753	1	US-08-712-241-3	Sequence 3, App11
588	6	0.6	652	3	US-08-993-170A-111	Sequence 111, App	661	6	0.6	753	4	US-08-026-143B-3	Sequence 3, App11
589	6	0.6	652	3	US-08-993-775B-2	Sequence 2, App11	662	6	0.6	753	5	PCT-US92-10621-3	Sequence 3, App11
590	6	0.6	652	3	US-08-993-775B-4	Sequence 4, App11	663	6	0.6	753	5	PCT-US94-02233-3	Sequence 3, App11
591	6	0.6	652	3	US-08-993-775B-6	Sequence 6, App11	664	6	0.6	753	5	PCT-US94-02233-3	Sequence 3, App11
592	6	0.6	652	3	US-08-993-775B-8	Sequence 8, App11	665	6	0.6	754	2	US-08-941-262-1	Sequence 1, App11
593	6	0.6	652	3	US-08-993-775B-10	Sequence 10, App1	666	6	0.6	756	4	US-09-085-199B-9	Sequence 9, App11
594	6	0.6	652	3	US-08-993-775B-12	Sequence 12, App1	667	6	0.6	775	2	US-08-714-070A-1	Sequence 1, App11
595	6	0.6	652	3	US-08-993-775B-14	Sequence 14, App1	668	6	0.6	790	4	US-09-723-153-2	Sequence 2, App11
596	6	0.6	652	3	US-08-993-775B-16	Sequence 16, App1	669	6	0.6	795	4	US-09-193-5620-11	Sequence 11, App1
597	6	0.6	652	3	US-08-993-775B-18	Sequence 18, App1	670	6	0.6	796	4	US-08-857-076-40	Sequence 40, App1
598	6	0.6	652	3	US-08-993-775B-20	Sequence 20, App1	671	6	0.6	798	4	US-08-867-076-40	Sequence 40, App1
599	6	0.6	652	3	US-08-993-775B-22	Sequence 22, App1	672	6	0.6	800	3	US-08-867-076-40	Sequence 40, App1
600	6	0.6	652	3	US-08-993-775B-24	Sequence 24, App1	673	6	0.6	804	4	US-08-867-076-40	Sequence 40, App1
601	6	0.6	652	3	US-08-993-775B-26	Sequence 26, App1	674	6	0.6	821	4	US-08-855-910-4	Sequence 4, App11
602	6	0.6	652	3	US-08-993-775B-28	Sequence 28, App1	675	6	0.6	833	2	US-09-193-5620-12	Sequence 12, App1
603	6	0.6	652	3	US-08-993-775B-30	Sequence 30, App1	676	6	0.6	833	2	US-08-844-086A-2	Sequence 2, App11
604	6	0.6	652	3	US-08-993-775B-32	Sequence 32, App1	677	6	0.6	835	2	US-09-018-211-2	Sequence 2, App11
605	6	0.6	652	3	US-08-993-775B-34	Sequence 34, App1	678	6	0.6	836	1	US-08-968-751-4	Sequence 4, App11
606	6	0.6	652	3	US-08-993-775B-36	Sequence 36, App1	679	6	0.6	836	1	US-07-998-003A-105	Sequence 105, App
607	6	0.6	652	3	US-08-993-775B-38	Sequence 38, App1	680	6	0.6	836	1	US-08-453-724B-105	Sequence 105, App
608	6	0.6	652	3	US-08-993-775B-40	Sequence 40, App1	681	6	0.6	836	1	US-08-453-695A-105	Sequence 105, App
609	6	0.6	652	3	US-08-993-775B-42	Sequence 42, App1	682	6	0.6	836	2	US-08-268-161A-105	Sequence 105, App
610	6	0.6	652	3	US-08-993-775B-44	Sequence 44, App1	683	6	0.6	836	4	US-08-453-702A-105	Sequence 105, App
611	6	0.6	652	3	US-08-993-775B-46	Sequence 46, App1	684	6	0.6	836	5	US-09-099-639-105	Sequence 105, App
												PCT-US93-12588-105	Sequence 105, App

685	6	0.6	836	5	PCT-US95-08071-105	Sequence 105, App	758	6	0.6	1013	4	US-09-240-473-5	Sequence 5, Appl
686	6	0.6	845	6	US-08-356-354-2	Patent No. 5196194	759	6	0.6	1038	4	US-09-541-782-4	Sequence 4, Appl
687	6	0.6	846	1	US-06-356-354-2	Sequence 2, Appl	760	6	0.6	1054	4	US-08-356-354-4	Sequence 4, Appl
688	6	0.6	846	2	US-07-728-215-33	Sequence 33, Appl	761	6	0.6	1054	2	US-08-778-656-4	Sequence 4, Appl
689	6	0.6	846	2	US-08-778-656-2	Sequence 2, Appl	762	6	0.6	1069	1	US-07-777-715-9	Sequence 9, Appl
690	6	0.6	846	4	US-08-938-085A-33	Sequence 33, Appl	763	6	0.6	1069	1	US-08-170-126-4	Sequence 4, Appl
691	6	0.6	858	4	US-08-857-076-41	Sequence 41, Appl	764	6	0.6	1069	1	US-08-954-418-4	Sequence 4, Appl
692	6	0.6	864	2	US-08-209-521-29	Sequence 29, Appl	765	6	0.6	1084	1	US-08-717-515-6	Sequence 6, Appl
693	6	0.6	876	1	US-08-717-515-4	Sequence 4, Appl	766	6	0.6	1090	3	US-08-307-896-3	Sequence 3, Appl
694	6	0.6	879	1	US-08-220-151-2	Sequence 2, Appl	767	6	0.6	1090	3	US-08-726-214-4	Sequence 4, Appl
695	6	0.6	879	1	US-08-220-151-3	Sequence 3, Appl	768	6	0.6	1090	4	US-09-085-199B-5	Sequence 5, Appl
696	6	0.6	879	1	US-08-413-118-2	Sequence 2, Appl	769	6	0.6	1090	4	PCT-US95-11808-3	Sequence 3, Appl
697	6	0.6	879	1	US-08-413-118-3	Sequence 3, Appl	770	6	0.6	1098	1	US-07-777-715-7	Sequence 7, Appl
698	6	0.6	879	3	US-08-413-118-106	Sequence 106, Appl	771	6	0.6	1098	1	US-08-170-126-2	Sequence 2, Appl
699	6	0.6	879	3	US-08-473-446-2	Sequence 2, Appl	772	6	0.6	1098	3	US-08-954-418-2	Sequence 2, Appl
700	6	0.6	879	3	US-08-473-446-3	Sequence 3, Appl	773	6	0.6	1123	4	US-09-408-865-1	Sequence 1, Appl
701	6	0.6	879	3	US-08-473-446-106	Sequence 106, Appl	774	6	0.6	1129	4	US-08-357-598-6	Sequence 6, Appl
702	6	0.6	892	4	US-08-857-076-42	Sequence 42, Appl	775	6	0.6	1129	1	US-08-097-997A-9	Sequence 9, Appl
703	6	0.6	894	1	US-08-117-362-4	Sequence 4, Appl	776	6	0.6	1129	2	US-08-567-508C-3	Sequence 3, Appl
704	6	0.6	894	1	US-08-486-924-4	Sequence 4, Appl	777	6	0.6	1129	2	US-09-003-289-6	Sequence 6, Appl
705	6	0.6	898	2	US-08-808-982-5	Sequence 5, Appl	778	6	0.6	1129	3	US-09-196-480-3	Sequence 9, Appl
706	6	0.6	898	4	US-09-306-902A-5	Sequence 5, Appl	779	6	0.6	1129	4	US-08-665-574C-9	Sequence 9, Appl
707	6	0.6	903	1	US-08-021-601-12	Sequence 12, Appl	780	6	0.6	1129	5	US-08-946-994-9	Sequence 9, Appl
708	6	0.6	903	1	US-08-082-849B-12	Sequence 12, Appl	781	6	0.6	1132	2	PCT-US95-16435-6	Sequence 6, Appl
709	6	0.6	903	4	US-09-193-562D-46	Sequence 46, Appl	782	6	0.6	1132	2	US-08-567-508C-2	Sequence 2, Appl
710	6	0.6	904	1	PCT-US94-01624-12	Sequence 12, Appl	783	6	0.6	1132	3	US-09-196-480-2	Sequence 2, Appl
711	6	0.6	904	1	US-07-998-003A-97	Sequence 97, Appl	784	6	0.6	1139	4	US-09-046-188A-22	Sequence 22, Appl
712	6	0.6	904	1	US-08-453-274B-97	Sequence 97, Appl	785	6	0.6	1163	1	US-08-173-497-4	Sequence 4, Appl
713	6	0.6	904	1	US-08-453-695A-97	Sequence 97, Appl	786	6	0.6	1163	1	US-08-286-889-4	Sequence 4, Appl
714	6	0.6	904	1	US-08-268-161A-97	Sequence 97, Appl	787	6	0.6	1163	1	US-08-486-618-4	Sequence 4, Appl
715	6	0.6	904	2	US-08-453-702A-97	Sequence 97, Appl	788	6	0.6	1163	1	US-08-362-652-4	Sequence 4, Appl
716	6	0.6	904	2	US-09-093-639-97	Sequence 97, Appl	789	6	0.6	1163	2	US-08-605-672-4	Sequence 4, Appl
717	6	0.6	904	5	PCT-US93-12588-97	Sequence 97, Appl	790	6	0.6	1163	2	US-08-482-293A-4	Sequence 4, Appl
718	6	0.6	904	5	PCT-US95-08071-97	Sequence 97, Appl	791	6	0.6	1163	2	US-08-943-363-4	Sequence 4, Appl
719	6	0.6	905	4	US-09-193-562D-2	Sequence 2, Appl	792	6	0.6	1163	2	US-08-476-062A-44	Sequence 44, Appl
720	6	0.6	910	4	US-08-460-269C-2	Sequence 4, Appl	793	6	0.6	1163	5	US-09-193-043-4	Sequence 5, Appl
721	6	0.6	914	4	US-09-085-199B-4	Sequence 4, Appl	794	6	0.6	1163	5	PCT-US96-01314-44	Sequence 44, Appl
722	6	0.6	921	1	US-08-188-582-2	Sequence 2, Appl	795	6	0.6	1182	4	US-09-041-886-21	Sequence 21, Appl
723	6	0.6	921	1	US-08-646-715-2	Sequence 2, Appl	796	6	0.6	1248	3	US-08-726-214-16	Sequence 16, Appl
724	6	0.6	924	3	US-08-619-812-8	Sequence 8, Appl	797	6	0.6	1276	1	US-08-717-515-8	Sequence 8, Appl
725	6	0.6	926	1	US-07-908-253-2	Sequence 2, Appl	798	6	0.6	1323	4	US-09-004-838-90	Sequence 90, Appl
726	6	0.6	926	1	US-08-455-970A-2	Sequence 2, Appl	799	6	0.6	1334	6	US-08-004-838-90	Sequence 90, Appl
727	6	0.6	926	1	US-08-387-156-6	Sequence 6, Appl	800	6	0.6	1363	1	US-08-425-061-23	Sequence 23, Appl
728	6	0.6	926	2	US-08-694-865-6	Sequence 6, Appl	801	6	0.6	1363	2	US-08-825-886-23	Sequence 23, Appl
729	6	0.6	926	2	US-08-878-748-6	Sequence 6, Appl	802	6	0.6	1382	3	US-09-057-570-4	Sequence 4, Appl
730	6	0.6	926	2	US-08-535-837-2	Sequence 2, Appl	803	6	0.6	1403	3	US-07-908-253-3	Sequence 3, Appl
731	6	0.6	926	3	US-09-124-491-6	Sequence 6, Appl	804	6	0.6	1403	2	US-08-694-865-17	Sequence 17, Appl
732	6	0.6	926	6	547657-3	Patent No. 547657	805	6	0.6	1403	2	US-08-535-837-3	Sequence 3, Appl
733	6	0.6	928	3	US-09-320-878-13	Sequence 13, Appl	806	6	0.6	1403	3	US-09-124-491-17	Sequence 17, Appl
734	6	0.6	928	4	US-09-105-537-41	Sequence 41, Appl	807	6	0.6	1449	3	US-08-840-062-6	Sequence 6, Appl
735	6	0.6	934	1	US-08-215-805A-80	Sequence 80, Appl	808	6	0.6	1507	6	5268270-2	Patent No. 5268270
736	6	0.6	936	1	US-08-455-970A-12	Sequence 12, Appl	809	6	0.6	1568	4	US-09-181-706-2	Sequence 2, Appl
737	6	0.6	943	1	US-08-455-970A-10	Sequence 10, Appl	810	6	0.6	1568	4	US-09-458-791-2	Sequence 2, Appl
738	6	0.6	944	4	US-09-449-285A-2	Sequence 2, Appl	811	6	0.6	1568	4	US-09-459-066-2	Sequence 2, Appl
739	6	0.6	951	1	US-08-445-970A-14	Sequence 14, Appl	812	6	0.6	1621	4	US-08-927-927-3	Sequence 3, Appl
740	6	0.6	971	1	US-08-446-038B-19	Sequence 19, Appl	813	6	0.6	1657	3	US-09-057-570-2	Sequence 2, Appl
741	6	0.6	971	1	US-08-446-010B-19	Sequence 19, Appl	814	6	0.6	1658	2	US-08-609-049A-13	Sequence 13, Appl
742	6	0.6	971	2	US-08-805-445-19	Sequence 19, Appl	815	6	0.6	1726	2	US-09-170-996A-30	Sequence 30, Appl
743	6	0.6	971	2	US-08-064-067D-19	Sequence 19, Appl	816	6	0.6	1726	2	US-08-609-049A-30	Sequence 30, Appl
744	6	0.6	971	2	US-09-066-208-19	Sequence 19, Appl	817	6	0.6	1754	1	US-09-170-996-30	Sequence 30, Appl
745	6	0.6	977	1	US-08-387-156-8	Sequence 8, Appl	818	6	0.6	1754	1	US-07-745-206A-13	Sequence 13, Appl
746	6	0.6	977	2	US-08-694-865-8	Sequence 8, Appl	819	6	0.6	1754	2	US-08-311-363-13	Sequence 13, Appl
747	6	0.6	977	2	US-08-878-748-8	Sequence 8, Appl	820	6	0.6	1786	4	US-08-444-818-54	Sequence 54, Appl
748	6	0.6	977	3	US-09-124-491-8	Sequence 8, Appl	821	6	0.6	1788	2	US-08-962-284-2	Sequence 2, Appl
749	6	0.6	979	3	US-08-870-529-2	Sequence 2, Appl	822	6	0.6	1792	2	US-08-962-284-4	Sequence 4, Appl
750	6	0.6	986	3	US-08-938-291A-7	Sequence 7, Appl	823	6	0.6	1805	3	US-09-057-570-7	Sequence 7, Appl
751	6	0.6	993	1	US-08-446-010B-25	Sequence 25, Appl	824	6	0.6	1817	4	US-09-004-838-125	Sequence 125, Appl
752	6	0.6	993	4	US-08-980-080-4	Sequence 4, Appl	825	6	0.6	1824	2	US-08-680-327-3	Sequence 3, Appl
753	6	0.6	1000	4	US-09-193-562D-30	Sequence 30, Appl	826	6	0.6	1824	4	US-09-228-246-2	Sequence 2, Appl
754	6	0.6	1003	2	US-07-743-357-9	Sequence 9, Appl	827	6	0.6	1852	1	US-08-425-061-24	Sequence 24, Appl
755	6	0.6	1013	2	US-08-866-650-5	Sequence 5, Appl	828	6	0.6	1852	2	US-08-825-886-24	Sequence 24, Appl
756	6	0.6	1013	3	US-09-021-287-5	Sequence 5, Appl	829	6	0.6	1863	1	US-08-425-061-16	Sequence 16, Appl
757	6	0.6	1013	3	US-08-991-408-2	Sequence 2, Appl	830	6	0.6	1863	1	US-08-598-591-2	Sequence 2, Appl

831	6	0.6	1863	1	US-08-480-784-2	Sequence 2, Appl1	904	6	0.6	3011	3	US-08-811-566-20	Sequence 20, Appl1
832	6	0.6	1863	1	US-08-483-553-2	Sequence 2, Appl1	905	6	0.6	3011	4	US-08-444-818-177	Sequence 177, App
833	6	0.6	1863	1	US-08-487-002-2	Sequence 2, Appl1	906	6	0.6	3011	4	US-09-014-416-1	Sequence 1, Appl1
834	6	0.6	1863	1	US-08-483-554B-2	Sequence 2, Appl1	907	6	0.6	3011	4	US-09-014-416-5	Sequence 5, Appl1
835	6	0.6	1863	1	US-08-798-691-2	Sequence 2, Appl1	908	6	0.6	3011	4	US-08-529-169A-6	Sequence 6, Appl1
836	6	0.6	1863	1	US-08-798-691-4	Sequence 4, Appl1	909	6	0.6	3011	4	US-09-388-874-2	Sequence 2, Appl1
837	6	0.6	1863	1	US-08-798-691-6	Sequence 6, Appl1	910	6	0.6	3011	4	US-09-046-604-36	Sequence 36, Appl1
838	6	0.6	1863	1	US-08-488-011B-2	Sequence 2, Appl1	911	6	0.6	3011	4	US-08-440-549-10	Sequence 10, Appl
839	6	0.6	1863	2	US-08-825-886-16	Sequence 16, Appl	912	6	0.6	3011	5	PCT-US91-32225-10	Sequence 10, Appl
840	6	0.6	1863	2	US-08-603-753D-2	Sequence 2, Appl1	913	6	0.6	3011	5	PCT-US93-00907-1	Sequence 1, Appl1
841	6	0.6	1863	3	US-08-825-487A-2	Sequence 2, Appl1	914	6	0.6	3011	5	PCT-US93-00907-2	Sequence 2, Appl1
842	6	0.6	1863	3	US-08-825-487A-4	Sequence 4, Appl1	915	6	0.6	3011	5	PCT-US94-07280-1	Sequence 1, Appl1
843	6	0.6	1863	3	US-08-825-487A-6	Sequence 6, Appl1	916	6	0.6	3011	5	PCT-US95-01087-1	Sequence 1, Appl1
844	6	0.6	1863	4	US-09-074-476-2	Sequence 2, Appl1	917	6	0.6	3012	3	US-08-811-566-2	Sequence 2, Appl1
845	6	0.6	1863	4	US-09-074-476-4	Sequence 4, Appl1	918	6	0.6	3457	2	US-08-416-603-4	Sequence 4, Appl1
846	6	0.6	1863	4	US-09-074-476-6	Sequence 6, Appl1	919	6	0.6	3567	2	US-07-642-734C-4	Sequence 4, Appl1
847	6	0.6	1863	4	US-09-099-753-2	Sequence 2, Appl1	920	6	0.6	3567	3	US-08-439-009A-4	Sequence 4, Appl1
848	6	0.6	1863	4	US-08-850-727-2	Sequence 2, Appl1	921	6	0.6	3788	4	US-09-336-447A-76	Sequence 76, Appl1
849	6	0.6	1863	4	US-08-986-106-2	Sequence 2, Appl1	922	6	0.6	4472	2	US-08-804-227C-2	Sequence 2, Appl1
850	6	0.6	1863	4	US-09-007-678B-49	Sequence 49, Appl1	923	6	0.6	4544	1	US-08-469-486-52	Sequence 52, Appl
851	6	0.6	1863	5	PCT-US95-10203-2	Sequence 2, Appl1	924	6	0.6	4544	2	US-08-469-486-52	Sequence 14, Appl
852	6	0.6	1863	5	PCT-US95-10203-2	Sequence 2, Appl1	925	6	0.6	4545	2	US-08-804-227C-14	Sequence 14, Appl
853	6	0.6	1863	5	PCT-US95-10220-2	Sequence 2, Appl1	926	6	0.6	4531	3	US-09-320-878-1	Sequence 1, Appl1
854	6	0.6	1898	2	US-08-056-200-94	Sequence 94, Appl	927	6	0.6	4613	4	US-09-105-537-31	Sequence 31, Appl
855	6	0.6	1898	2	US-08-800-644-94	Sequence 94, Appl	928	6	0.6	5087	4	US-09-144-085-1	Sequence 1, Appl1
856	6	0.6	1940	2	US-08-644-271-30	Sequence 30, Appl	929	6	0.6	5588	4	US-09-036-987A-6	Sequence 6, Appl1
857	6	0.6	2030	1	US-07-551-531-2	Sequence 2, Appl1	930	6	0.6	5588	4	US-09-370-700-6	Sequence 6, Appl1
858	6	0.6	2109	4	US-08-646-695-6	Sequence 6, Appl1	931	6	0.6	6095	4	US-09-144-085-2	Sequence 2, Appl1
859	6	0.6	2109	4	PCT-US96-06053-6	Sequence 6, Appl1	932	6	0.6	11877	4	US-09-105-537-6	Sequence 6, Appl1
860	6	0.6	2115	4	US-09-324-867-5	Sequence 6, Appl1	933	6	0.5	5	1	US-08-097-938-40	Sequence 40, Appl
861	6	0.6	2133	2	US-08-670-707A-37	Sequence 37, Appl	934	5	0.5	5	1	US-08-476-000-40	Sequence 40, Appl
862	6	0.6	2133	2	US-09-037-601-37	Sequence 37, Appl	935	5	0.5	5	1	US-08-472-840-40	Sequence 40, Appl
863	6	0.6	2206	1	US-07-852-260-2	Sequence 2, Appl1	936	5	0.5	5	2	US-08-476-976-40	Sequence 40, Appl
864	6	0.6	2206	2	US-08-461-503-2	Sequence 2, Appl1	937	5	0.5	5	3	US-08-476-976-40	Sequence 40, Appl
865	6	0.6	2206	4	US-08-465-250-2	Sequence 2, Appl1	938	5	0.5	5	4	US-08-486-673B-40	Sequence 40, Appl
866	6	0.6	2237	1	US-08-354-973-1	Sequence 1, Appl1	939	5	0.5	6	1	US-08-061-350-6	Sequence 6, Appl1
867	6	0.6	2261	4	US-08-444-818-66	Sequence 66, Appl	940	5	0.5	6	1	US-08-097-938-23	Sequence 23, Appl
868	6	0.6	2265	2	US-08-149-097D-36	Sequence 36, Appl	941	5	0.5	6	1	US-08-097-938-24	Sequence 24, Appl
869	6	0.6	2411	4	US-09-268-347-36	Sequence 36, Appl	942	5	0.5	6	1	US-08-476-000-39	Sequence 39, Appl
870	6	0.6	2436	4	US-08-444-818-75	Sequence 75, Appl	943	5	0.5	6	1	US-08-097-938-55	Sequence 55, Appl
871	6	0.6	2485	5	PCT-US94-00198-1	Sequence 1, Appl1	944	5	0.5	6	1	US-08-476-000-23	Sequence 23, Appl
872	6	0.6	2485	5	PCT-US94-00198-2	Sequence 2, Appl1	945	5	0.5	6	1	US-08-476-000-24	Sequence 24, Appl
873	6	0.6	2509	2	US-08-149-097D-35	Sequence 35, Appl	946	5	0.5	6	1	US-08-476-000-39	Sequence 39, Appl
874	6	0.6	2539	4	US-09-413-814-42	Sequence 42, Appl	947	5	0.5	6	1	US-08-476-000-55	Sequence 55, Appl
875	6	0.6	2618	4	US-09-413-814-28	Sequence 28, Appl	948	5	0.5	6	1	US-08-472-840-23	Sequence 23, Appl
876	6	0.6	2772	4	US-08-444-818-89	Sequence 89, Appl	949	5	0.5	6	1	US-08-472-840-24	Sequence 24, Appl
877	6	0.6	2818	1	US-08-510-284-1	Sequence 1, Appl1	950	5	0.5	6	1	US-08-472-840-39	Sequence 39, Appl
878	6	0.6	2818	1	US-08-411-389-2	Sequence 2, Appl1	951	5	0.5	6	1	US-08-472-840-55	Sequence 55, Appl
879	6	0.6	2818	4	US-07-966-049A-2	Sequence 2, Appl1	952	5	0.5	6	2	US-08-323-436A-44	Sequence 44, Appl
880	6	0.6	2818	4	US-09-542-331-2	Sequence 2, Appl1	953	5	0.5	6	2	US-08-413-490-44	Sequence 44, Appl
881	6	0.6	2884	2	US-08-466-975A-23	Sequence 23, Appl	954	5	0.5	6	2	US-08-476-976-23	Sequence 23, Appl
882	6	0.6	2884	2	US-08-391-671A-23	Sequence 23, Appl	955	5	0.5	6	2	US-08-476-976-39	Sequence 39, Appl
883	6	0.6	2884	2	US-08-467-902A-23	Sequence 23, Appl	956	5	0.5	6	2	US-08-476-976-39	Sequence 39, Appl
884	6	0.6	2894	3	US-09-275-265-23	Sequence 23, Appl	957	5	0.5	6	2	US-08-476-976-55	Sequence 55, Appl
885	6	0.6	2894	3	US-08-443-260-3	Sequence 3, Appl1	958	5	0.5	6	2	US-08-476-976-55	Sequence 55, Appl
886	6	0.6	2955	2	US-08-442-805A-3	Sequence 3, Appl1	959	5	0.5	6	3	US-08-474-410-23	Sequence 23, Appl
887	6	0.6	2955	3	US-08-443-900A-3	Sequence 3, Appl1	960	5	0.5	6	3	US-08-474-410-39	Sequence 39, Appl
888	6	0.6	2955	4	US-08-444-818-124	Sequence 124, App	961	5	0.5	6	3	US-08-474-410-39	Sequence 39, Appl
889	6	0.6	2955	4	US-08-249-843-3	Sequence 3, Appl1	962	5	0.5	6	3	US-09-131-956-1	Sequence 55, Appl
890	6	0.6	2955	4	US-08-444-818-138	Sequence 138, App	963	5	0.5	6	3	US-09-146-269-44	Sequence 44, Appl
891	6	0.6	3011	1	US-08-188-281B-1	Sequence 1, Appl1	964	5	0.5	6	3	US-09-146-269-44	Sequence 44, Appl
892	6	0.6	3011	1	US-08-453-552-1	Sequence 1, Appl1	965	5	0.5	6	4	US-08-486-673B-23	Sequence 23, Appl
893	6	0.6	3011	1	US-08-453-552-2	Sequence 2, Appl1	966	5	0.5	6	4	US-08-486-673B-23	Sequence 23, Appl
894	6	0.6	3011	1	US-08-440-103-36	Sequence 36, Appl	967	5	0.5	6	4	US-08-486-673B-39	Sequence 39, Appl
895	6	0.6	3011	1	US-08-440-542-36	Sequence 36, Appl	968	5	0.5	6	4	US-08-486-673B-39	Sequence 39, Appl
896	6	0.6	3011	1	US-07-910-760-10	Sequence 10, Appl	969	5	0.5	6	4	US-09-294-457-1	Sequence 55, Appl
897	6	0.6	3011	1	US-08-440-519-10	Sequence 10, Appl	970	5	0.5	6	4	PCT-US95-13037-44	Sequence 44, Appl
898	6	0.6	3011	1	US-08-231-368-36	Sequence 36, Appl	971	5	0.5	6	5	5194592-19	Sequence 1, Appl1
899	6	0.6	3011	1	US-08-440-210-36	Sequence 36, Appl	972	5	0.5	6	6	US-08-097-938-38	Sequence 38, Appl
900	6	0.6	3011	2	US-08-710-637-1	Sequence 1, Appl1	973	5	0.5	7	1	US-08-476-000-38	Sequence 38, Appl
901	6	0.6	3011	2	US-08-710-637-2	Sequence 2, Appl1	974	5	0.5	7	1	US-08-472-840-38	Sequence 38, Appl
902	6	0.6	3011	2	US-08-833-678A-6	Sequence 6, Appl1	975	5	0.5	7	1	US-08-472-840-38	Sequence 38, Appl
903	6	0.6	3011	2	US-08-833-678A-6	Sequence 6, Appl1	976	5	0.5	7	2	US-08-476-976-38	Sequence 38, Appl

```

977 5 0.5 7 2 US-08-819-013-3 Sequence 3, Appli
978 5 0.5 7 3 US-08-474-410-38 Sequence 38, Appl
979 5 0.5 7 4 US-08-486-673B-38 Sequence 38, Appl
980 5 0.5 8 1 US-08-190-788A-19 Sequence 19, Appl
981 5 0.5 8 1 US-08-097-938-37 Sequence 37, Appl
982 5 0.5 8 1 US-08-476-000-37 Sequence 37, Appl
983 5 0.5 8 1 US-08-472-840-37 Sequence 37, Appl
984 5 0.5 8 1 US-08-383-474B-24 Sequence 24, Appl
985 5 0.5 8 1 US-08-465-391A-19 Sequence 19, Appl
986 5 0.5 8 2 US-08-464-538B-19 Sequence 19, Appl
987 5 0.5 8 2 US-08-476-976-37 Sequence 37, Appl
988 5 0.5 8 2 US-08-463-076E-63 Sequence 63, Appl
989 5 0.5 8 2 US-08-350-260A-436 Sequence 436, Appl
990 5 0.5 8 3 US-08-159-339A-1046 Sequence 1046, Ap
991 5 0.5 8 3 US-08-474-410-37 Sequence 37, Appl
992 5 0.5 8 4 US-08-486-673B-37 Sequence 37, Appl
993 5 0.5 9 1 US-08-288-405A-6 Sequence 6, Appl
994 5 0.5 9 1 US-08-097-938-36 Sequence 36, Appl
995 5 0.5 9 1 US-08-476-000-36 Sequence 36, Appl
996 5 0.5 9 1 US-08-472-840-36 Sequence 36, Appl
997 5 0.5 9 2 US-08-650-262-12 Sequence 12, Appl
998 5 0.5 9 2 US-08-476-976-36 Sequence 36, Appl
999 5 0.5 9 3 US-08-159-339A-742 Sequence 742, App
1000 5 0.5 9 3 US-08-159-339A-1054 Sequence 1054, Ap

```

ALIGNMENTS

```

RESULT 1
US-08-679-493A-185
; Sequence 185, Application US/08679493A
; Patent No. 6303295
; GENERAL INFORMATION:
; APPLICANT: Taylor, Ethan W.
; TITLE OF INVENTION: SELENOPROTEINS, CODING SEQUENCES AND METHODS
; FILE REFERENCE: 55-95
; CURRENT APPLICATION NUMBER: US/08/679,493A
; PRIOR FILING DATE: 1996-07-12
; PRIOR APPLICATION NUMBER: 60/001203
; PRIOR FILING DATE: 1995-07-14
; PRIOR APPLICATION NUMBER: 60/003,112
; PRIOR FILING DATE: 1995-09-01
; NUMBER OF SEQ ID NOS: 216
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 185
; LENGTH: 196
; TYPE: PRT
; ORGANISM: Tetrahymena sp.
US-08-679-493A-185

```

```

Query Match          0.7%; Score 7; DB 4; Length 196;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1060 EKKRYLOA 1066
DB 188 EKKRYLOA 194

```

```

RESULT 2
US-08-459-264-2
; Sequence 2, Application US/08459264
; Patent No. 5705340
; GENERAL INFORMATION:
; APPLICANT: RASMUSSEN, Beth A
; APPLICANT: TALLY, Francis P
; APPLICANT: GUZMAN, Yakov
; TITLE OF INVENTION: DNA SEQUENCES AND AMINO ACID SEQUENCES
; TITLE OF INVENTION: OF CLASS B BETA-LACTAMASE ENZYMES FROM BACTEROIDES
; TITLE OF INVENTION: FRAGILIS
; NUMBER OF SEQUENCES: 9

```

```

CORRESPONDENCE ADDRESS:
ADDRESS: Darby & Darby PC
STREET: 805 Third Avenue
CITY: New York
STATE: NY
COUNTRY: US
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/459,264
FILING DATE: 02-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Robinson, Joseph R
REGISTRATION NUMBER: 33,448
REFERENCE/DOCKET NUMBER: 0646/1B026-US3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-527-7700
TELEFAX: 212-753-6237
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 249 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Bacteroides fragilis
IMMEDIATE SOURCE:
CLONE: corr amino acid
US-08-459-264-2

```

```

Query Match          0.7%; Score 7; DB 1; Length 249;
Best Local Similarity 100.0%; Pred. No. 86;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 619 DKVYTYV 625
DB 35 DKVYTYV 41

```

```

RESULT 3
US-08-459-264-3
; Sequence 3, Application US/08459264
; Patent No. 5705340
; GENERAL INFORMATION:
; APPLICANT: RASMUSSEN, Beth A
; APPLICANT: TALLY, Francis P
; APPLICANT: GUZMAN, Yakov
; TITLE OF INVENTION: DNA SEQUENCES AND AMINO ACID SEQUENCES
; TITLE OF INVENTION: OF CLASS B BETA-LACTAMASE ENZYMES FROM BACTEROIDES
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Darby & Darby PC
; STREET: 805 Third Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/459,264

```

FILING DATE: 02-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: ROBINSON, JOSEPH R
REGISTRATION NUMBER: 33,448
REFERENCE/DOCKET NUMBER: 0646/1B026-US3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-527-7700
TELEFAX: 212-753-6237
TELEX: 236687
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 249 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Bacteroides fragilis
IMMEDIATE SOURCE:
CLONE: tal3635
US-08-459-264-3

Query Match 0.7%; Score 7; DB 1; Length 249;
Best Local Similarity 100.0%; Pred. No. 86;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 619 DKVYTYV 625
|||||
DB 35 DKVYTYV 41

RESULT 4
US-08-459-263-2
Sequence 2, Application US/08459263
Patent No. 5792642
GENERAL INFORMATION:
APPLICANT: RASMUSSEN, Beth A
APPLICANT: TALLY, Francis P
APPLICANT: GLUZMAN, Yakov
TITLE OF INVENTION: DNA SEQUENCES AND AMINO ACID SEQUENCES
TITLE OF INVENTION: OF CLASS B BETA-LACTAMASE ENZYMES FROM BACTEROIDES
TITLE OF INVENTION: FRAGILIS
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Darby & Darby PC
STREET: 805 Third Avenue
CITY: New York
STATE: NY
COUNTRY: US
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/459,263
FILING DATE: 02-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: ROBINSON, JOSEPH R
REGISTRATION NUMBER: 33,448
REFERENCE/DOCKET NUMBER: 0646/1B026-US2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-527-7700
TELEFAX: 212-753-6237
TELEX: 236687
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 249 amino acids
TYPE: amino acid

STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Bacteroides fragilis
IMMEDIATE SOURCE:
CLONE: ccra amino acid
US-08-459-263-2

Query Match 0.7%; Score 7; DB 1; Length 249;
Best Local Similarity 100.0%; Pred. No. 86;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 619 DKVYTYV 625
|||||
DB 35 DKVYTYV 41

RESULT 5
US-08-459-263-3
Sequence 3, Application US/08459263
Patent No. 5792642
GENERAL INFORMATION:
APPLICANT: RASMUSSEN, Beth A
APPLICANT: TALLY, Francis P
APPLICANT: GLUZMAN, Yakov
TITLE OF INVENTION: DNA SEQUENCES AND AMINO ACID SEQUENCES
TITLE OF INVENTION: OF CLASS B BETA-LACTAMASE ENZYMES FROM BACTEROIDES
TITLE OF INVENTION: FRAGILIS
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Darby & Darby PC
STREET: 805 Third Avenue
CITY: New York
STATE: NY
COUNTRY: US
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/459,263
FILING DATE: 02-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: ROBINSON, JOSEPH R
REGISTRATION NUMBER: 33,448
REFERENCE/DOCKET NUMBER: 0646/1B026-US2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-527-7700
TELEFAX: 212-753-6237
TELEX: 236687
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 249 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Bacteroides fragilis
IMMEDIATE SOURCE:
CLONE: tal3635
US-08-459-263-3

Query Match 0.7%; Score 7; DB 1; Length 249;
Best Local Similarity 100.0%; Pred. No. 86;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 619 DKYTYV 625
DB 35 DKYTYV 41

RESULT 6

US-08-809-103B-2
; Sequence 2, Application US/08809103B
; Patent No. 6133505
; GENERAL INFORMATION:
; APPLICANT: GRONENBORN, Bruno
; TITLE OF INVENTION: PHYTOPATHOGENIC DNA VIRUS RESISTANT
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: YOUNG & THOMPSON
; STREET: 745 South 23rd Street
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/809,103B
; FILING DATE: 17-MAR-1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 94.11040
; FILING DATE: 15-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/FR95/01192
; FILING DATE: 15-SEP-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: PATCH, Andrew J.
; REGISTRATION NUMBER: 32,925
; REFERENCE/DOCKET NUMBER: USB94AL CNR TOM
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 521-2297
; TELEFAX: (703) 685-0573
; TELEX: 248425 EMBON
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 359 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-809-103B-2

Query Match 0.7%; Score 7; DB 4; Length 359;

Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 103 DKDGYL 109
DB 105 DKDGYL 111

RESULT 7

US-08-809-103B-4
; Sequence 4, Application US/08809103B
; Patent No. 6133505
; GENERAL INFORMATION:
; APPLICANT: GRONENBORN, Bruno
; TITLE OF INVENTION: PHYTOPATHOGENIC DNA VIRUS RESISTANT
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: YOUNG & THOMPSON

STREET: 745 South 23rd Street
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/809,103B
FILING DATE: 17-MAR-1997
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 94.11040
FILING DATE: 15-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/FR95/01192
FILING DATE: 15-SEP-1995
ATTORNEY/AGENT INFORMATION:
NAME: PATCH, Andrew J.
REGISTRATION NUMBER: 32,925
REFERENCE/DOCKET NUMBER: USB94AL CNR TOM
TELEPHONE: (703) 521-2297
TELEFAX: (703) 685-0573
TELEX: 248425 EMBON
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 359 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-809-103B-4

Query Match 0.7%; Score 7; DB 4; Length 359;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 103 DKDGYL 109
DB 105 DKDGYL 111

RESULT 8

US-08-809-103B-6
; Sequence 6, Application US/08809103B
; Patent No. 6133505
; GENERAL INFORMATION:
; APPLICANT: GRONENBORN, Bruno
; TITLE OF INVENTION: PHYTOPATHOGENIC DNA VIRUS RESISTANT
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: YOUNG & THOMPSON
; STREET: 745 South 23rd Street
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/809,103B
; FILING DATE: 17-MAR-1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 94.11040

FILING DATE: 15-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/FR95/01192
FILING DATE: 15-SEP-1995
ATTORNEY/AGENT INFORMATION:
NAME: PATCH, Andrew J.
REGISTRATION NUMBER: 32,925
REFERENCE/DOCKET NUMBER: USB94AL CNR TOM
TELEPHONE: (703) 521-2297
TELEFAX: (703) 685-0573
TELEX: 248425 EMBON
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 359 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-809-103b-6

Query Match 0.7%; Score 7; DB 4; Length 359;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 103 DKDGDVL 109
|||||
DB 105 DKDGDVL 111

RESULT 9
US-08-809-103b-8
Sequence 8, Application US/08809103B
Patent No. 6133505
GENERAL INFORMATION:
APPLICANT: GROENBORN, Bruno
TITLE OF INVENTION: PHYTOPATHOGENIC DNA VIRUS RESISTANT
TITLE OF INVENTION: TRANSGENIC PLANTS AND SEEDS AND METHODS FOR OBTAINING SAME
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: YOUNG & THOMPSON
STREET: 745 South 23rd Street
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/809,103B
FILING DATE: 17-MAR-1997
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 94.11040
FILING DATE: 15-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/FR95/01192
FILING DATE: 15-SEP-1995
ATTORNEY/AGENT INFORMATION:
NAME: PATCH, Andrew J.
REGISTRATION NUMBER: 32,925
REFERENCE/DOCKET NUMBER: USB94AL CNR TOM
TELEPHONE: (703) 521-2297
TELEFAX: (703) 685-0573
TELEX: 248425 EMBON
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 359 amino acids
TYPE: amino acid

TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-809-103b-8

Query Match 0.7%; Score 7; DB 4; Length 359;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 103 DKDGDVL 109
|||||
DB 105 DKDGDVL 111

RESULT 10
US-08-118-270-15
Sequence 15, Application US/08118270
Patent No. 5508384
GENERAL INFORMATION:
APPLICANT: Murphy, Randall B.
TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN
TITLE OF INVENTION: RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF
NUMBER OF SEQUENCES: 348
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEIMARK
STREET: 419 Seventh Street, N.W., Suite 300
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/118,270
FILING DATE: 09-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/943,236
FILING DATE: 10-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: Townsend, Kevin G.
REGISTRATION NUMBER: 34,033
REFERENCE/DOCKET NUMBER: MURPHY-2A
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
TELEX: 248633
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 362 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-118-270-15

Query Match 0.7%; Score 7; DB 1; Length 362;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 744 VGIWWSL 750
|||||
DB 2 VGIWWSL 8

RESULT 11
PCT-US93-08528-15
Sequence 15, Application PC/TUS9308528
GENERAL INFORMATION:

APPLICANT: New York University
TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN
TITLE OF INVENTION: RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF
NUMBER OF SEQUENCES: 348
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROADY AND NEWMARK
STREET: 419 Seventh Street, N.W., Suite 300
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/08528
FILING DATE: 09-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/943,236
FILING DATE: 10-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: Townsend, Kevin G.
REGISTRATION NUMBER: 34,033
REFERENCE/DOCKET NUMBER: MURPHY-2 PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
TELEX: 248633
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 362 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
PCT-US93-08528-15

Query Match 0.7%; Score 7; DB 5; Length 362;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 744 VGIWMSL 750
DB 2 VGIWMSL 8

RESULT 12
US-09-021-323-1
Sequence 1, Application US/09021323
Patent No. 5929033
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Corley, Neil C.
APPLICANT: Yue, Henry
TITLE OF INVENTION: EXTRACELLULAR MUCOUS MATRIX
TITLE OF INVENTION: GLYCOPROTEIN
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Dr.
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/021,323
FILING DATE: Filed Herewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0477 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 510 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: SINTNOT13
CLONE: 1805538
US-09-021-323-1

Query Match 0.7%; Score 7; DB 2; Length 510;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 SSRSQSS 38
DB 49 SSRSQSS 55

RESULT 13
US-08-907-166-12
Sequence 12, Application US/08907166
Patent No. 5948666
GENERAL INFORMATION:
APPLICANT: Callen, Walter
APPLICANT: Mather, Eric
TITLE OF INVENTION: ISOLATION AND IDENTIFICATION OF NOVEL POLYMERASES
FILE REFERENCE: 09010/027001
CURRENT APPLICATION NUMBER: US/08/907,166
CURRENT FILING DATE: 1997-08-06
NUMBER OF SEQ ID NOS: 12
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 12
LENGTH: 574
TYPE: PRT
ORGANISM: Aquifex pyrophillus
US-08-907-166-12

Query Match 0.7%; Score 7; DB 2; Length 574;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 243 SVVLGKK 249
DB 387 SVVLGKK 393

RESULT 14
US-08-420-235B-21
Sequence 21, Application US/08420235B
Patent No. 5801042
GENERAL INFORMATION:
APPLICANT: Chang, Yuan
APPLICANT: Moore, Patrick S.
TITLE OF INVENTION: UNIQUE ASSOCIATED KAPOSI'S SARCOMA VIRUS
TITLE OF INVENTION: SEQUENCES AND USES THEREOF
NUMBER OF SEQUENCES: 47

Job time: 184 sec

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; City: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/420,235B
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 45185-B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 752 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-420-235B-21

Query Match 0.7%: Score 7; DB 1; Length 752;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 43 RIFSLLE 49
|||||||
Db 602 RIFSLLE 608

RESULT 15
US-08-793-624-21
; Sequence 21, Application US/08793624C
; Patent No. 6150093
; GENERAL INFORMATION:
; APPLICANT: Chang, Yuan
; APPLICANT: Moore, Patrick S.
; TITLE OF INVENTION: Unique Associated Kaposi's Sarcoma Virus Sequences And
; FILE REFERENCE: 45185-C-PCT-US/JPM
; CURRENT APPLICATION NUMBER: US/08/793,624C
; CURRENT FILING DATE: 1997-02-18
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 21
; LENGTH: 752
; TYPE: PRT
; ORGANISM: Kaposi's sarcoma-associated herpesvirus
; US-08-793-624-21

Query Match 0.7%: Score 7; DB 4; Length 752;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 43 RIFSLLE 49
|||||||
Db 602 RIFSLLE 608

Search completed: May 21, 2002, 14:54:19

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 21, 2002, 14:52:01 ; Search time 26.02 Seconds
(without alignments)
3966.178 Million cell updates/sec

Title: US-09-729-653-2_COPY_1_1074
Perfect score: 1074
Sequence: 1 HSLIGRCSRGLGDGNVAC.....LYFEGEKRYLQAGKFFLLCG 1074

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 283138 seqs, 96089334 residues

Word size: 0

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	9	0.8	519	2 T27880	hypothetical prote
2	8	0.7	176	2 T17821	hypothetical prote
3	8	0.7	241	2 A84283	hypothetical prote
4	8	0.7	316	2 AE3648	flagellar motor sw
5	8	0.7	394	1 KIBSGM	phosphoglycerate k
6	8	0.7	414	2 AC0575	hypothetical membr
7	8	0.7	519	2 E87233	probable secreted
8	8	0.7	642	2 T29967	hypothetical prote
9	8	0.7	843	2 T34618	NADH dehydrogenase
10	8	0.7	1214	2 T47659	spliceosomal-like
11	8	0.7	1238	2 T32625	hypothetical prote
12	8	0.7	1520	2 G98275	hypothetical prote
13	8	0.7	1520	2 AF3008	polyketide synthas
14	7	0.7	39	2 A96026	probable transpos
15	7	0.7	91	4 S14968	hypothetical NADH
16	7	0.7	94	2 F84497	En/spm-like transp
17	7	0.7	98	2 S26981	pvs protein 1 - Ki
18	7	0.7	100	2 D48223	Ig heavy chain V r
19	7	0.7	110	2 A95296	hypothetical prote
20	7	0.7	123	2 G86281	protein F10B6.16 l
21	7	0.7	130	2 AC2314	hypothetical prote
22	7	0.7	131	1 S30119	ribosomal protein
23	7	0.7	136	2 G64734	yael protein - Esc
24	7	0.7	136	2 C90644	hypothetical prote
25	7	0.7	136	2 C85495	hypothetical prote
26	7	0.7	142	2 G87147	conserved hypochet
27	7	0.7	143	2 T40376	heat shock protein
28	7	0.7	144	1 R5EC15	ribosomal protein
29	7	0.7	144	2 F64094	ribosomal protein

30	7	0.7	144	2 F91149	50S ribosomal subu
31	7	0.7	144	2 B85995	50S ribosomal subu
32	7	0.7	144	2 AD1008	50S ribosomal chat
33	7	0.7	154	2 G86699	hypothetical prote
34	7	0.7	154	2 J10071	alpha-pilin - Mora
35	7	0.7	155	2 A41490	beta precursor -
36	7	0.7	156	2 A24434	beta fibrillar prot
37	7	0.7	157	2 AB2539	hypothetical prote
38	7	0.7	158	2 A55851	prepilin - Moraxel
39	7	0.7	160	2 A25571	myosin LC1 catalyt
40	7	0.7	164	2 A42460	fibrinai protein Q
41	7	0.7	173	2 S23411	FUN3 protein - ye
42	7	0.7	183	1 RDQFBR	ubiquinol--cytochr
43	7	0.7	196	2 A39223	superoxide dismuta
44	7	0.7	201	2 E70211	hypothetical prote
45	7	0.7	201	2 F95878	probable ISRM2011-
46	7	0.7	216	2 S61416	dihydrokaempferol
47	7	0.7	219	1 F69500	riboflavin-specific
48	7	0.7	219	2 A26484	glutathione transf
49	7	0.7	223	2 D70760	hypothetical prote
50	7	0.7	227	2 B48536	DNA topoisomerase
51	7	0.7	229	2 D82230	hypothetical prote
52	7	0.7	232	2 C84706	probable small nuc
53	7	0.7	248	2 C75140	hypothetical prote
54	7	0.7	249	2 A35263	beta-lactamase (EC
55	7	0.7	250	2 A82125	heme exporter prot
56	7	0.7	259	2 T29727	hypothetical prote
57	7	0.7	262	2 F84994	glutamate racemase
58	7	0.7	265	2 S19113	cger-4 protein - C
59	7	0.7	267	2 C64175	hypothetical prote
60	7	0.7	275	2 B64986	hypothetical 30.9
61	7	0.7	275	2 G91011	probable elongatio
62	7	0.7	275	2 A85856	probable elongatio
63	7	0.7	282	2 B98166	hypothetical prote
64	7	0.7	282	2 A99205	hypothetical prote
65	7	0.7	282	2 AE3121	hypothetical prote
66	7	0.7	293	2 T47708	hypothetical prote
67	7	0.7	295	2 A49906	melibiose-1-phosphat
68	7	0.7	302	1 RGEGB	regulator of melib
69	7	0.7	302	2 D91266	regulator of melib
70	7	0.7	302	2 A86107	regulator of melib
71	7	0.7	305	2 JC4525	nucleic acid-bindi
72	7	0.7	307	2 A10004	conserved membrane
73	7	0.7	310	2 H81161	cysteine synthase
74	7	0.7	310	2 A11022	melibiose operon r
75	7	0.7	312	2 E70376	exopolysphatase
76	7	0.7	314	2 H71802	probable outer mem
77	7	0.7	314	2 T31220	transposase homolo
78	7	0.7	315	2 C64716	rare lipoprotein A
79	7	0.7	315	2 T25297	hypothetical prote
80	7	0.7	330	2 T32080	hypothetical prote
81	7	0.7	332	2 S64375	probable tyrosine
82	7	0.7	333	2 AH0268	anthranilate phosp
83	7	0.7	337	1 DEJUGC	glyceraldenhyde-3-p
84	7	0.7	338	2 T47218	glyceraldenhyde-3-p
85	7	0.7	338	2 D87801	protein C10G11.7 l
86	7	0.7	340	2 E95844	probable sugar ABC
87	7	0.7	342	2 AB2397	ATP-binding protei
88	7	0.7	344	2 G70036	spore coat polysac
89	7	0.7	350	2 S47292	phenol 2-monooxyge
90	7	0.7	351	2 S50098	site-specific DNA-
91	7	0.7	352	2 T22464	hypothetical prote
92	7	0.7	353	1 A4308	phenol 2-monooxyge
93	7	0.7	353	2 S70629	Rad51 homolog mei-
94	7	0.7	356	2 A99901	D-alanine-D-alanin
95	7	0.7	359	2 S22593	hypothetical prote
96	7	0.7	359	2 S39211	gene C1 protein -
97	7	0.7	359	2 S39235	gene C1 protein -
98	7	0.7	364	1 MENZNC	matrix protein - N
99	7	0.7	364	1 MENZNV	matrix protein - N
100	7	0.7	365	1 B69114	conserved hypochet
101	7	0.7	370	1 W2WL39	E2 protein - human
102	7	0.7	371	2 A71683	hypothetical prote

50S ribosomal subu	50S ribosomal chat	hypothetical prote	alpha-pilin - Mora	beta precursor -	beta fibrillar prot	hypothetical prote	prepilin - Moraxel	myosin LC1 catalyt	fibrinai protein Q	FUN3 protein - ye	ubiquinol--cytochr	superoxide dismuta	hypothetical prote	probable ISRM2011-	dihydrokaempferol	riboflavin-specific	glutathione transf	hypothetical prote	DNA topoisomerase	hypothetical prote	probable small nuc	hypothetical prote	beta-lactamase (EC	heme exporter prot	hypothetical prote	glutamate racemase	cger-4 protein - C	hypothetical prote	hypothetical 30.9	probable elongatio	probable elongatio	hypothetical prote	hypothetical prote	hypothetical prote	melibiose-1-phosphat	regulator of melib	regulator of melib	nucleic acid-bindi	conserved membrane	cysteine synthase	melibiose operon r	exopolysphatase	probable outer mem	transposase homolo	rare lipoprotein A	hypothetical prote	hypothetical prote	probable tyrosine	anthranilate phosp	glyceraldenhyde-3-p	glyceraldenhyde-3-p	protein C10G11.7 l	probable sugar ABC	ATP-binding protei	spore coat polysac	phenol 2-monooxyge	site-specific DNA-	hypothetical prote	phenol 2-monooxyge	Rad51 homolog mei-	D-alanine-D-alanin	hypothetical prote	gene C1 protein -	gene C1 protein -	matrix protein - N	matrix protein - N	conserved hypochet	E2 protein - human	hypothetical prote
--------------------	--------------------	--------------------	--------------------	------------------	---------------------	--------------------	--------------------	--------------------	--------------------	-------------------	--------------------	--------------------	--------------------	--------------------	-------------------	---------------------	--------------------	--------------------	-------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	-------------------	--------------------	--------------------	--------------------	--------------------	--------------------	----------------------	--------------------	--------------------	--------------------	--------------------	-------------------	--------------------	-----------------	--------------------	--------------------	--------------------	--------------------	--------------------	-------------------	--------------------	---------------------	---------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	-------------------	-------------------	--------------------	--------------------	--------------------	--------------------	--------------------

103	7	0.7	379	2	G70918	hypothetical prote	176	7	0.7	906	2	G70767	probable helicase
104	7	0.7	392	2	D71175	hypothetical prote	177	7	0.7	943	2	F69543	ATP-dependent RNA
105	7	0.7	394	2	JQ1399	phosphoglycerate k	178	7	0.7	957	2	A49847	nitrite reductase
106	7	0.7	394	2	A72223	pantothenate metab	179	7	0.7	986	2	JC4825	glucan 1,4-beta-gl
107	7	0.7	396	2	E90238	threonine synthase	180	7	0.7	988	1	DJVZFP	DNA-directed DNA p
108	7	0.7	396	2	F75388	1-deoxy-D-xylulose	181	7	0.7	993	2	S48436	probable RNA helic
109	7	0.7	400	2	AD2322	phosphoglycerate k	182	7	0.7	1056	2	T02930	lysine-ketoglutarat
110	7	0.7	401	1	TWWTGY	phosphoglycerate k	183	7	0.7	1062	2	F69102	protein F25E5.1 (I
111	7	0.7	402	2	S75082	pantothenate metab	184	7	0.7	1063	2	S18211	hypothetical prote
112	7	0.7	406	2	T36632	probable oxidoredu	185	7	0.7	1099	2	T30307	rexS protein - Lac
113	7	0.7	415	2	T40535	probable arginine-	186	7	0.7	1151	2	T18297	zinc-finger protei
114	7	0.7	418	2	AC3519	isochorismate synt	187	7	0.7	1159	2	B98198	hypothetical prote
115	7	0.7	425	2	F70608	hypothetical prote	188	7	0.7	1159	2	AH3088	Icmf (Imported) -
116	7	0.7	433	2	D84667	probable PCI domai	189	7	0.7	1188	2	F64367	pyruvate, water dik
117	7	0.7	435	2	A72658	probable Isocitrat	190	7	0.7	1297	2	S39791	neurotoxin - Clost
118	7	0.7	440	2	E71299	conserved hypothet	191	7	0.7	1334	2	T41524	rho1 gdp-gtp excha
119	7	0.7	450	2	S73419	signal recognition	192	7	0.7	1403	2	S24548	homeoic protein p
120	7	0.7	460	2	G90890	hypothetical prote	193	7	0.7	1474	2	T20488	hypothetical prote
121	7	0.7	460	2	A86180	hypothetical prote	194	7	0.7	1612	2	S59699	DNA topoisomerase
122	7	0.7	460	2	E84902	hypothetical prote	195	7	0.7	1626	2	A39242	DNA topoisomerase
123	7	0.7	460	2	B85727	hypothetical prote	196	7	0.7	1643	2	D71630	outer membrane pro
124	7	0.7	462	2	T07989	hypothetical prote	197	7	0.7	1732	2	T43026	probable DNA-direc
125	7	0.7	462	2	T52110	acetyl-CoA C-acylt	198	7	0.7	2090	2	S26058	trichocyst protein
126	7	0.7	468	2	A13241	conserved hypothet	199	7	0.7	2149	2	C96695	t-complex polypept
127	7	0.7	473	2	D81322	outer membrane pro	200	7	0.7	2391	2	G89779	glucagon G1 - Nort
128	7	0.7	481	2	B70908	hypothetical prote	201	7	0.7	2351	2	B98047	glucagon G2 - Nort
129	7	0.7	485	1	S32433	methyiaspartate mu	202	7	0.7	3305	2	T18358	hypothetical prote
130	7	0.7	488	2	F97039	hypothetical prote	203	6	0.6	15	2	S29174	apolipoprotein prec
131	7	0.7	489	2	T29924	hypothetical prote	204	6	0.6	28	2	A60698	D-galactose-bindin
132	7	0.7	498	2	T49747	hypothetical prote	205	6	0.6	29	2	A49410	trichocyst protein
133	7	0.7	510	2	A87482	conserved hypothet	206	6	0.6	31	2	S44471	t-complex polypept
134	7	0.7	515	2	F70904	hypothetical prote	207	6	0.6	31	2	S44472	glucagon G2 - Nort
135	7	0.7	517	2	T19962	hypothetical prote	208	6	0.6	31	2	C97398	hypothetical prote
136	7	0.7	520	2	D90014	hypothetical prote	209	6	0.6	32	2	S58524	alpha-complex prot
137	7	0.7	522	2	A83804	involved in spore	210	6	0.6	39	2	J00282	hypothetical 4.3K
138	7	0.7	531	2	T04463	hypothetical prote	211	6	0.6	44	2	F97562	hypothetical prote
139	7	0.7	535	2	B4576	D2 protein precurs	212	6	0.6	49	2	G85815	hypothetical prote
140	7	0.7	545	2	B82740	two-component syst	213	6	0.6	52	1	R3KM72	ribosomal protein
141	7	0.7	547	2	JQ0356	cellulase (EC 3.2.	214	6	0.6	53	2	T09190	hypothetical prote
142	7	0.7	551	2	S57447	HPBR11-7 protein -	215	6	0.6	53	2	T28821	hypothetical prote
143	7	0.7	556	2	S39732	arginine/ornithine	216	6	0.6	54	2	H90803	hypothetical prote
144	7	0.7	573	2	H86313	protein F2H15.10 (217	6	0.6	56	2	E69938	hypothetical prote
145	7	0.7	574	2	G70468	DNA polymerase I (218	6	0.6	57	2	S63680	signal transducer
146	7	0.7	576	2	A40688	peroxisomal protei	219	6	0.6	57	2	T10457	Lipoprotein L - Ps
147	7	0.7	589	2	H69691	two-component sens	220	6	0.6	61	2	A81898	hypothetical prote
148	7	0.7	591	2	A75474	serine proteinase,	221	6	0.6	63	2	S08190	metallothionein 1
149	7	0.7	595	2	T39681	transcription regu	222	6	0.6	64	2	T35034	hypothetical prote
150	7	0.7	596	2	A90534	hypothetical prote	223	6	0.6	65	2	D97706	hypothetical prote
151	7	0.7	605	2	D71053	hypothetical prote	224	6	0.6	66	2	S16201	photosystem I chai
152	7	0.7	607	2	A55898	probable sensor Ki	225	6	0.6	68	2	E70002	conserved hypothet
153	7	0.7	608	2	H75089	twlching mobility	226	6	0.6	68	2	F95222	hypothetical prote
154	7	0.7	612	2	I64241	glucose inhibited	227	6	0.6	68	2	F98086	hypothetical prote
155	7	0.7	625	2	T41603	alpha-amylase - fi	228	6	0.6	69	2	B91001	hypothetical prote
156	7	0.7	629	2	A95948	probable chaperoni	229	6	0.6	70	2	S35030	photosystem I chai
157	7	0.7	632	2	E81684	1-deoxyxylulose-5-	230	6	0.6	70	2	T06922	photosystem I chai
158	7	0.7	640	2	F71527	probable transkeo	231	6	0.6	71	2	T60082	CDA receptor - hum
159	7	0.7	654	2	G72344	phosphoglycerate k	232	6	0.6	71	2	H95299	hypothetical prote
160	7	0.7	664	2	AF1600	transketolase homo	233	6	0.6	72	2	E90412	hypothetical prote
161	7	0.7	678	2	T05821	hypothetical prote	234	6	0.6	76	2	G83935	thioredoxin relate
162	7	0.7	694	2	C72761	hypothetical prote	235	6	0.6	77	2	D82819	hypothetical prote
163	7	0.7	695	2	D71283	probable translati	236	6	0.6	78	2	A62080	hypothetical prote
164	7	0.7	752	2	T16508	hypothetical prote	237	6	0.6	79	2	B64614	hypothetical prote
165	7	0.7	758	2	T16800	hypothetical prote	238	6	0.6	79	2	C71901	hypothetical prote
166	7	0.7	769	2	A71403	probable glucan sy	239	6	0.6	80	2	T28256	ORF MSV095 hypothe
167	7	0.7	784	2	E84785	probable protein k	240	6	0.6	81	2	S01410	hypothetical prote
168	7	0.7	803	2	H71437	probable protein k	241	6	0.6	82	2	D84562	hypothetical prote
169	7	0.7	807	2	T40821	probable ubiquitin	242	6	0.6	82	2	H83584	probable biotin-re
170	7	0.7	816	2	S54518	probable membrane	243	6	0.6	82	2	AD3480	hypothetical prote
171	7	0.7	825	2	H82885	hypothetical prote	244	6	0.6	84	2	A82916	ribosomal protein
172	7	0.7	846	2	C82135	chitinase VC1952 (245	6	0.6	84	2	C95923	hypothetical hypot
173	7	0.7	858	2	B86416	unknown protein, 1	246	6	0.6	85	2	T39205	dyein light chain
174	7	0.7	858	2	C86416	unknown protein, 2	247	6	0.6	87	2	B69420	hypothetical prote
175	7	0.7	889	2	B55123	coatmer complex b	248	6	0.6	88	2	S68159	H+-transporting AT

249	6	0.6	88	2	C85866	hypothetical prote	322	6	0.6	125	2	T03283	hypothetical prote
250	6	0.6	89	2	F69858	hypothetical prote	323	6	0.6	125	2	S23541	hypothetical prote
251	6	0.6	89	2	AF3734	conserved hypot	324	6	0.6	125	2	JC4799	basic leucine zipp
252	6	0.6	89	2	AG3394	hypothetical membr	325	6	0.6	125	2	AE1047	conserved hypot
253	6	0.6	90	2	C97737	hypothetical prote	326	6	0.6	126	2	C70307	ribosomal protein
254	6	0.6	91	2	H82370	conserved hypot	327	6	0.6	126	2	T40732	probable 50s ribs
255	6	0.6	91	2	B82709	hypothetical prote	328	6	0.6	126	2	AB2905	conserved hypot
256	6	0.6	92	2	S37507	Ig kappa chain V r	329	6	0.6	126	2	AG1932	hypothetical prote
257	6	0.6	92	2	S21307	hypothetical prote	330	6	0.6	126	2	T47476	hypothetical prote
258	6	0.6	94	2	AG3478	hypothetical prote	331	6	0.6	127	2	PC2266	cytochrome P450 pr
259	6	0.6	95	2	D69121	hypothetical prote	332	6	0.6	127	2	T20707	hypothetical prote
260	6	0.6	97	2	H70197	stage V sporulatio	333	6	0.6	127	2	H64012	sufti protein homol
261	6	0.6	98	2	C91022	polymyxin resist	334	6	0.6	128	2	S27919	hypothetical prote
262	6	0.6	98	2	A64997	polymyxin B resist	335	6	0.6	128	2	J02182	hypothetical 14.8K
263	6	0.6	98	2	D85866	polymyxin resist	336	6	0.6	128	2	G81035	hypothetical prote
264	6	0.6	98	2	A12262	hypothetical prote	337	6	0.6	128	2	D64251	RNaseP C5 chain -
265	6	0.6	99	2	E82930	conserved hypot	338	6	0.6	128	2	AD1648	hypothetical prote
266	6	0.6	100	2	E43550	hypothetical prote	339	6	0.6	129	2	C69488	LSU ribosomal prot
267	6	0.6	101	2	C90075	intercellular adhe	340	6	0.6	129	2	S63743	pol polyprotein -
268	6	0.6	101	2	D72318	hypothetical prote	341	6	0.6	129	2	D90367	hypothetical prote
269	6	0.6	102	2	S17687	rRNA N-glycosidase	342	6	0.6	129	2	G90022	hypothetical prote
270	6	0.6	102	2	S17688	rRNA N-glycosidase	343	6	0.6	130	2	G70799	hypothetical prote
271	6	0.6	102	2	S11894	rRNA N-glycosidase	344	6	0.6	130	2	AD2901	hypothetical prote
272	6	0.6	102	2	S76695	hypothetical prote	345	6	0.6	131	2	RDEB15	hypothetical prote
273	6	0.6	103	2	B87261	hypothetical prote	346	6	0.6	131	2	S14390	hypothetical prote
274	6	0.6	104	2	AB1214	TN916 ORF23 homolo	347	6	0.6	131	2	J02171	hypothetical prote
275	6	0.6	105	2	A48943	phase antigenic de	348	6	0.6	131	2	S27911	hypothetical prote
276	6	0.6	106	2	T49742	hypothetical prote	349	6	0.6	131	2	AB1525	hypothetical prote
277	6	0.6	107	2	A61431	peptidylprolyl iso	350	6	0.6	132	1	F71177	hypothetical prote
278	6	0.6	107	2	T29406	hypothetical prote	351	6	0.6	132	2	E75164	hypothetical prote
279	6	0.6	107	2	T49594	hypothetical prote	352	6	0.6	132	2	B71954	hypothetical prote
280	6	0.6	108	1	K3HUB6	Ig kappa chain V-I	353	6	0.6	132	2	B64554	conserved hypot
281	6	0.6	108	2	A35780	peptidylprolyl iso	354	6	0.6	132	2	T46905	hypothetical prote
282	6	0.6	108	2	UC5764	FK506-binding pro	355	6	0.6	132	2	G72537	hypothetical prote
283	6	0.6	108	2	A42657	FK506-binding pro	356	6	0.6	133	2	S58123	chlorodoxin (clone
284	6	0.6	108	2	JH0528	FK506-binding pro	357	6	0.6	133	2	T08447	hypothetical prote
285	6	0.6	108	2	T26539	hypothetical prote	358	6	0.6	133	2	T30655	hypothetical prote
286	6	0.6	108	2	E30609	Ig kappa chain V-I	359	6	0.6	133	2	D97680	hypothetical prote
287	6	0.6	108	2	F64010	hypothetical prote	360	6	0.6	133	2	A41328	hypothetical prote
288	6	0.6	111	2	S74191	ribonuclease (EC 3	361	6	0.6	133	2	G70727	hypothetical prote
289	6	0.6	111	2	S11978	white protein VII -	362	6	0.6	134	1	G65129	hypothetical prote
290	6	0.6	111	2	H72707	hypothetical prote	363	6	0.6	134	2	AD2592	cytidine deaminase
291	6	0.6	112	2	AB3507	phosphoribosyl-Atp	364	6	0.6	134	2	G95280	probable regulator
292	6	0.6	112	2	D75451	hypothetical prote	365	6	0.6	134	2	G91154	hypothetical prote
293	6	0.6	114	2	T23119	hypothetical prote	366	6	0.6	134	2	D86000	hypothetical prote
294	6	0.6	114	2	G97938	hypothetical prote	367	6	0.6	134	2	AB1002	conserved hypot
295	6	0.6	115	1	A46279	guanylin precursor	368	6	0.6	134	2	AC1166	hypothetical prote
296	6	0.6	115	2	S03243	hypothetical prote	369	6	0.6	134	2	H90235	hypothetical prote
297	6	0.6	116	1	S20630	somatosatin precu	370	6	0.6	134	2	B86199	hypothetical prote
298	6	0.6	116	2	T49862	related to REM1 pr	371	6	0.6	135	2	S00388	T-cell receptor ga
299	6	0.6	117	2	S34246	hypothetical prote	372	6	0.6	135	2	AB3150	hypothetical prote
300	6	0.6	117	2	T17315	hypothetical prote	373	6	0.6	135	2	AC1447	hypothetical prote
301	6	0.6	119	2	T19321	hypothetical prote	374	6	0.6	135	2	AD1361	Porlein gpII (Bact
302	6	0.6	119	2	T29306	hypothetical prote	375	6	0.6	136	2	G97606	hypothetical prote
303	6	0.6	119	2	A85816	unknown protein en	376	6	0.6	136	2	AI2828	conserved hypot
304	6	0.6	119	2	B75331	conserved hypot	377	6	0.6	137	2	E69359	hypothetical prote
305	6	0.6	120	2	C91027	D-erythro-7,8-di	378	6	0.6	137	2	C75600	DNA-binding protel
306	6	0.6	120	2	D85871	D-erythro-7,8-di	379	6	0.6	138	2	AC3337	hypothetical prote
307	6	0.6	120	2	E65002	D-erythro-7,8-di	380	6	0.6	139	2	H72016	ribonuclease p pro
308	6	0.6	120	2	F81701	hypothetical prote	381	6	0.6	139	2	D86607	ribonuclease p pro
309	6	0.6	120	2	H97515	hypothetical prote	382	6	0.6	139	2	G90466	4-carboxymuolacto
310	6	0.6	122	2	T02926	acyl carrier prote	383	6	0.6	139	2	C87459	hypothetical prote
311	6	0.6	121	2	G64315	hypothetical prote	384	6	0.6	139	2	PC4217	hypothetical 139 p
312	6	0.6	122	2	D83448	succinate dehydro	385	6	0.6	139	2	B95304	hypothetical prote
313	6	0.6	122	2	E83111	transcription fact	386	6	0.6	139	2	AB2127	hypothetical prote
314	6	0.6	122	2	A41810	hypothetical prote	387	6	0.6	140	2	F72263	flagellar basal-do
315	6	0.6	122	2	B75387	hypothetical prote	388	6	0.6	140	2	T10925	3C3, 16c protein -
316	6	0.6	123	2	H90672	hypothetical prote	389	6	0.6	140	2	T46220	hypothetical prote
317	6	0.6	123	2	B84752	hypothetical prote	390	6	0.6	140	2	A11741	protein involved i
318	6	0.6	124	1	C08Y26	cell division cont	391	6	0.6	140	2	G86413	F28N24, 17 protein
319	6	0.6	124	1	XVVCB	cholera enterotoxi	392	6	0.6	141	1	HA1PI	hemoglobin alpha-I
320	6	0.6	124	1	F72615	hypothetical prote	393	6	0.6	141	1	HA1P2B	hemoglobin alpha-I
321	6	0.6	124	2	B82444	hypothetical prote	394	6	0.6	141	1	H69385	hypothetical prote

395	6	0.6	141	2	S55247	hemoglobin alpha-A
396	6	0.6	141	2	D64530	ribosomal protein
397	6	0.6	141	2	T1975	ribosomal protein
398	6	0.6	141	2	T40501	60s ribosomal prot
399	6	0.6	141	2	D69857	conserved hypotnet
400	6	0.6	141	2	D82939	small protein B uu
401	6	0.6	141	2	E70869	hypothetical prote
402	6	0.6	141	2	H64361	hypothetical prote
403	6	0.6	141	2	AC1372	proteins involved
404	6	0.6	142	2	JH0586	ribosomal protein
405	6	0.6	142	2	B70634	probable mmpsi pro
406	6	0.6	142	2	A95132	hypothetical prote
407	6	0.6	142	2	AP1808	hypothetical prote
408	6	0.6	143	2	G96072	7,8-dihydro-8-oxog
409	6	0.6	143	2	S73515	PTS system mannito
410	6	0.6	144	2	F84988	50S ribosomal prot
411	6	0.6	144	2	S43766	ribosomal protein
412	6	0.6	144	2	T43774	ribosomal protein
413	6	0.6	144	2	T64121	DNA-directed DNA p
414	6	0.6	144	2	A84263	hypothetical prote
415	6	0.6	144	2	AE3423	hypothetical prote
416	6	0.6	144	2	G69255	hypothetical prote
417	6	0.6	144	2	D96967	probable metal-dep
418	6	0.6	145	2	F97374	Cytidine deaminase
419	6	0.6	145	2	A87618	C-factor imported
420	6	0.6	145	2	E84938	flagellar filip pro
421	6	0.6	146	1	G69110	conserved hypotnet
422	6	0.6	146	2	F70313	hypothetical prote
423	6	0.6	146	2	T48566	hemoglobin epsilon
424	6	0.6	147	1	HECH	ribosomal protein
425	6	0.6	147	2	T35563	hypothetical prote
426	6	0.6	147	2	D98205	hypothetical prote
427	6	0.6	147	2	AD3081	conserved hypotnet
428	6	0.6	147	2	C95227	conserved hypotnet
429	6	0.6	147	2	G98091	conserved hypotnet
430	6	0.6	147	2	D69544	hypothetical prote
431	6	0.6	147	2	S50775	hypothetical prote
432	6	0.6	147	2	H82172	heat shock protein
433	6	0.6	147	2	I40475	hypothetical prote
434	6	0.6	147	2	S72994	hypothetical prote
435	6	0.6	148	2	AE0928	probable phage tai
436	6	0.6	148	2	A11058	probable membrane
437	6	0.6	149	2	F75609	response regulator
438	6	0.6	149	2	T37099	hypothetical prote
439	6	0.6	150	2	T14457	MAUS box protein h
440	6	0.6	150	2	H83632	conserved hypotnet
441	6	0.6	151	1	S73506	PIIB homolog K05.o
442	6	0.6	151	2	I38365	epididymal secreto
443	6	0.6	151	2	I53929	epididymal secreto
444	6	0.6	152	2	G95958	probable transcrip
445	6	0.6	153	1	WRBP15	early protein gp17
446	6	0.6	153	2	AF1334	hypothetical prote
447	6	0.6	154	2	A44809	fimbrial protein -
448	6	0.6	154	2	D82723	transcription elon
449	6	0.6	155	2	S32194	ribosomal protein
450	6	0.6	155	2	T49854	hypothetical prote
451	6	0.6	155	2	D87414	hypothetical prote
452	6	0.6	156	2	G82952	ATP synthase B cha
453	6	0.6	156	2	T43957	hypothetical prote
454	6	0.6	156	2	C97097	probable membrane
455	6	0.6	157	2	H64376	phosphoribosylamin
456	6	0.6	157	2	S17934	tRNA N-glycosidase
457	6	0.6	157	2	T10732	intracellular path
458	6	0.6	158	2	T12958	copper transport p
459	6	0.6	158	2	D85523	hypothetical prote
460	6	0.6	159	2	A47699	type 4 pilin - Bix
461	6	0.6	159	2	G82924	hypothetical un180
462	6	0.6	159	2	H81657	conserved hypotnet
463	6	0.6	159	2	B84052	hypothetical prote
464	6	0.6	160	2	S40063	fimbrial protein c
465	6	0.6	160	2	S74317	fimbrial protein c
466	6	0.6	160	2	AT1562	molymbdenum cofacto
467	6	0.6	160	2	AF1205	molymbdenum cofacto
468	6	0.6	160	2	C70947	hypothetical prote
469	6	0.6	160	2	AH2342	hypothetical prote
470	6	0.6	161	2	T14817	pathogenesis-relat
471	6	0.6	161	2	D97145	molymbdenum cofacto
472	6	0.6	162	1	S07661	probable phosphor
473	6	0.6	162	2	C33739	hypothetical 17.8k
474	6	0.6	163	2	T11552	thioredoxin peroxi
475	6	0.6	163	2	B43583	conserved hypotnet
476	6	0.6	163	2	D75504	hypothetical prote
477	6	0.6	164	2	T35813	hypothetical prote
478	6	0.6	165	2	S35195	hypothetical prote
479	6	0.6	165	2	G84027	molymbdenum cofacto
480	6	0.6	165	2	T36924	probable alanine-r
481	6	0.6	165	2	T52114	probable transcrip
482	6	0.6	165	2	JE0065	retroviral protein
483	6	0.6	166	2	F97676	succinate dehydrog
484	6	0.6	166	2	H84952	hydroxanthine phosph
485	6	0.6	166	2	JN0030	early protein gp17
486	6	0.6	166	2	H87652	chemotaxis protein
487	6	0.6	166	2	T02222	NBS-LRR type resis
488	6	0.6	167	1	S48484	signal peptidase (
489	6	0.6	167	2	B85437	caltractin-like pr
490	6	0.6	167	2	B53293	flm3 region hypoth
491	6	0.6	167	2	S40978	hypothetical prote
492	6	0.6	168	2	H90399	hypothetical prote
493	6	0.6	169	2	T04207	phospholipid-hydro
494	6	0.6	169	2	F72465	hypothetical prote
495	6	0.6	170	2	B99552	adenine phosphorib
496	6	0.6	170	2	H69786	molymbdenum cofacto
497	6	0.6	170	2	G69541	conserved hypotnet
498	6	0.6	170	2	AC1219	diol dehydrase (di
499	6	0.6	170	2	AF1572	diol dehydrase (di
500	6	0.6	170	2	T48974	hypothetical prote
501	6	0.6	171	2	G87218	probable cytidine/
502	6	0.6	171	2	E83684	hypothetical prote
503	6	0.6	171	2	S15911	hypothetical prote
504	6	0.6	172	2	T32120	hypothetical prote
505	6	0.6	173	2	H82149	crossover junction
506	6	0.6	174	1	WRBP70	early protein gp17
507	6	0.6	174	2	E75552	ribosomal protein
508	6	0.6	174	2	T01649	probable trypsin 1
509	6	0.6	176	2	E90160	conserved hypotnet
510	6	0.6	176	2	C68824	hypothetical prote
511	6	0.6	176	2	C55208	socA3 protein - My
512	6	0.6	177	2	S36914	ribosomal protein
513	6	0.6	177	2	T37444	probable 20.7k pro
514	6	0.6	177	2	T40970	very hypothetical
515	6	0.6	178	2	S39875	hypothetical prote
516	6	0.6	178	2	G83844	hypothetical prote
517	6	0.6	178	2	T26853	hypothetical prote
518	6	0.6	179	2	AD1391	H+-transporting AT
519	6	0.6	179	2	B84008	hypothetical prote
520	6	0.6	179	2	AB1850	hypothetical prote
521	6	0.6	179	2	AF1766	H+-transporting AT
522	6	0.6	180	2	H81343	molymbdopterin bios
523	6	0.6	180	2	T18161	hypothetical prote
524	6	0.6	180	2	UC2394	W protein - Newcas
525	6	0.6	180	2	A60670	aducanin - mouse (i
526	6	0.6	180	2	T13878	matuase-like prot
527	6	0.6	181	2	T04398	hypothetical prote
528	6	0.6	181	2	AC2124	hypothetical prote
529	6	0.6	182	2	J01801	B7R 21.3k protein
530	6	0.6	182	2	G75343	conserved hypotnet
531	6	0.6	183	1	C70008	pyrazinamidase/nic
532	6	0.6	183	2	S42547	glucose-1-phosphat
533	6	0.6	183	2	G89872	conserved hypotnet
534	6	0.6	184	2	S75762	hypothetical prote
535	6	0.6	184	2	A37201	aminoglycoside N6'
536	6	0.6	184	2	E71002	hypothetical prote
537	6	0.6	185	2	JK0021	somatotropin - ski
538	6	0.6	185	2	A97118	probable sigma fac
539	6	0.6	186	2	JC2400	PMS5 homolog misma
540	6	0.6	187	2	PC1082	somatotropin - Jap

541	6	0.6	187	2	JU0030	somatotropin - yel	614	6	0.6	206	1	I64088	conserved hypotet
542	6	0.6	187	2	PC1088	somatotropin - gre	615	6	0.6	206	2	F87175	pyridoxamine 5'-ph
543	6	0.6	187	2	PC1087	somatotropin - dus	616	6	0.6	206	4	G47070	hypothetical tyrel
544	6	0.6	187	2	F64834	fimbrial-like prot	617	6	0.6	207	2	B70447	riboflavin synthas
545	6	0.6	187	2	C90757	probable fimbrial-	618	6	0.6	207	2	H84273	hypothetical Al precu
546	6	0.6	187	2	AH2190	polypeptide deform	619	6	0.6	207	2	F83815	hypothetical prote
547	6	0.6	188	2	F64496	GMP synthetase (EC	620	6	0.6	207	2	D86225	hypothetical prote
548	6	0.6	188	2	C87216	probable hemagglut	621	6	0.6	207	2	F83148	hypothetical prote
549	6	0.6	188	2	A84214	adenine phosphorib	622	6	0.6	207	2	T39349	probable 60S ribos
550	6	0.6	188	2	T15651	hypothetical prote	623	6	0.6	209	2	A71313	probable V-type At
551	6	0.6	189	2	T43766	hypothetical prote	624	6	0.6	209	2	S16864	hypothetical prote
552	6	0.6	190	2	AH0156	probable elongatio	625	6	0.6	209	2	AF2037	hypothetical prote
553	6	0.6	191	2	S09635	paba protein - Ser	626	6	0.6	209	2	B81809	hypothetical prote
554	6	0.6	191	2	AF0021	para-aminobenzoate	627	6	0.6	209	2	G82359	conserved hypotet
555	6	0.6	191	2	T07994	ribosomal protein	628	6	0.6	210	2	B84775	probable harpin-in
556	6	0.6	191	2	E64909	probable phage-rel	629	6	0.6	210	2	C97157	stage III sporulat
557	6	0.6	191	2	H64887	probable phage-rel	630	6	0.6	210	2	S64376	hypothetical prote
558	6	0.6	192	1	E64087	lipoprotein B - Ha	631	6	0.6	210	2	E84499	hypothetical prote
559	6	0.6	193	2	G64187	anthranilate synth	632	6	0.6	210	2	E87395	hypothetical prote
560	6	0.6	193	2	AD1949	hypothetical prote	633	6	0.6	210	2	S67771	endoplasmic reticu
561	6	0.6	193	2	F81779	probable transpos	634	6	0.6	211	2	E70045	two-component resp
562	6	0.6	194	1	QX8P3L	hypothetical prote	635	6	0.6	211	2	C83360	hypothetical prote
563	6	0.6	194	1	G64026	lacyl-carrier-prot	636	6	0.6	212	2	A81190	conserved hypotet
564	6	0.6	195	2	E27733	nitO protein - Azo	637	6	0.6	213	2	E95232	ABC transporter, A
565	6	0.6	195	2	A90467	hypothetical prote	638	6	0.6	213	2	G98096	hypothetical prote
566	6	0.6	196	1	HSX15A	histone H5A - Afri	639	6	0.6	214	2	T45714	endochitinase-like
567	6	0.6	196	2	A30484	histone H5B - Afri	640	6	0.6	214	2	G83488	probable permease
568	6	0.6	196	2	S06555	finger protein (cl	641	6	0.6	214	2	C71120	hypothetical prote
569	6	0.6	197	1	NMPS2P	anthranilate synth	642	6	0.6	214	2	C72372	pyrazinamidase/nic
570	6	0.6	197	2	T50025	hypothetical prote	643	6	0.6	214	2	S57813	hypothetical prote
571	6	0.6	197	2	C86748	hypothetical prote	644	6	0.6	214	2	T27259	hypothetical prote
572	6	0.6	197	2	H86741	conserved hypotet	645	6	0.6	214	2	A90028	urase accessory p
573	6	0.6	197	2	H97145	hypothetical prote	646	6	0.6	215	1	D71038	hypothetical prote
574	6	0.6	198	2	AB0112	probable 5-formyl	647	6	0.6	215	2	E96493	probable transpos
575	6	0.6	198	2	B89939	hypothetical prote	648	6	0.6	215	2	A83023	probable transcrip
576	6	0.6	198	2	C70570	hypothetical prote	649	6	0.6	216	2	T35881	thymidine kinase (
577	6	0.6	198	2	T13215	hypothetical prote	650	6	0.6	216	2	JE0297	DRE/CRT-binding pr
578	6	0.6	198	2	B84870	probable molybdopt	651	6	0.6	216	2	T51830	transcription fact
579	6	0.6	199	2	F70742	hypothetical prote	652	6	0.6	216	2	S48486	probable membrane
580	6	0.6	199	2	C75213	hypothetical prote	653	6	0.6	217	2	T42053	red2 protein - Str
581	6	0.6	200	1	H70409	conserved hypotet	654	6	0.6	217	2	F72293	5-methylthiodenos
582	6	0.6	200	2	A87671	cytochrome c oxida	655	6	0.6	217	2	A72024	conserved hypotet
583	6	0.6	200	2	T35699	probable sigma fac	656	6	0.6	217	2	A86600	CT734 hypothetical
584	6	0.6	200	2	T08966	hypothetical prote	657	6	0.6	217	2	G84380	hypothetical prote
585	6	0.6	200	2	B90300	hypothetical prote	658	6	0.6	217	2	T14778	hypothetical prote
586	6	0.6	200	2	AF1495	probable sugar-pho	659	6	0.6	218	2	JC1520	site-specific DNA
587	6	0.6	201	2	B88431	protein M88.2 (Imp	660	6	0.6	218	2	S75269	signal peptidase I
588	6	0.6	201	2	A87680	RNA polymerase sig	661	6	0.6	218	2	G82786	conserved hypotet
589	6	0.6	201	2	B72739	hypothetical prote	662	6	0.6	218	2	B82804	hypothetical prote
590	6	0.6	201	2	A13328	hypothetical prote	663	6	0.6	218	2	G72754	probable phosphoen
591	6	0.6	202	2	AG3507	imidazoteglycerol-	664	6	0.6	219	2	T45352	pyridoxamine 5'-ph
592	6	0.6	202	2	G84502	hypothetical prote	665	6	0.6	219	2	S72721	amidotransferase h
593	6	0.6	202	2	AB1441	probable orotate p	666	6	0.6	219	2	AF0639	flagellar basal bo
594	6	0.6	202	2	F97342	transcription regu	667	6	0.6	219	2	T38129	c-8 sterol isomera
595	6	0.6	202	2	E95883	probable transcrip	668	6	0.6	220	1	ISPUCH	chalcone isomerase
596	6	0.6	203	2	S00747	somatotropin precu	669	6	0.6	220	1	S01765	GTP-binding protei
597	6	0.6	203	2	C71375	conserved hypotet	670	6	0.6	220	2	D70350	dethiobiotin synth
598	6	0.6	203	2	H70336	hypothetical prote	671	6	0.6	220	2	H69257	hypothetical prote
599	6	0.6	204	1	STFL	somatotropin precu	672	6	0.6	221	2	T35154	hypothetical prote
600	6	0.6	204	2	S22616	hypothetical prote	673	6	0.6	221	2	S73349	adhesin PI precurs
601	6	0.6	204	2	S01746	somatotropin precu	674	6	0.6	221	2	AF3195	transcription regu
602	6	0.6	204	2	JH0577	somatotropin precu	675	6	0.6	222	1	R5MX1	ribosomal protein
603	6	0.6	204	2	S30491	somatotropin - Aca	676	6	0.6	222	2	B64895	hypothetical prote
604	6	0.6	204	2	A56904	somatotropin precu	677	6	0.6	222	2	D90883	hypothetical prote
605	6	0.6	204	2	JC4261	somatotropin precu	678	6	0.6	222	2	C85735	hypothetical prote
606	6	0.6	204	2	AF0914	conserved hypotet	679	6	0.6	222	2	AE0668	probable lipoprote
607	6	0.6	204	2	T03962	r40q3 protein - ri	680	6	0.6	222	2	S77526	hypothetical prote
608	6	0.6	204	2	S67295	probable membrane	681	6	0.6	222	2	C75539	conserved hypotet
609	6	0.6	205	2	S30739	hypothetical prote	682	6	0.6	223	1	S72693	dethiobiotin synth
610	6	0.6	205	2	H91325	hypothetical prote	683	6	0.6	223	1	YLHMP	serum amyloid P-co
611	6	0.6	205	2	H86072	hypothetical prote	684	6	0.6	223	1	B86968	conserved hypotet
612	6	0.6	205	2	H83689	hypothetical prote	685	6	0.6	223	2	T26763	hypothetical prote
613	6	0.6	205	2	A97161	cell division prot	686	6	0.6	223	2	T35908	probable membrane

687	6	0.6	223	2	S06576	finger protein (cl
688	6	0.6	223	2	AH0409	conserved hypothet
689	6	0.6	224	2	E83859	hypothetical prote
690	6	0.6	224	2	H71203	hypothetical prote
691	6	0.6	224	2	G75218	hypothetical prote
692	6	0.6	224	2	A64418	hypothetical prote
693	6	0.6	224	2	T51638	probable transcrip
694	6	0.6	225	2	G64702	hypothetical prote
695	6	0.6	225	2	F71817	hypothetical prote
696	6	0.6	226	2	D87061	dehiodiolin synth
697	6	0.6	226	2	C82124	oxidoreductase, sh
698	6	0.6	227	2	D97786	biotin synthetis p
699	6	0.6	227	2	AB1076	probable fibrillar
700	6	0.6	227	2	A71224	probable fibrillar
701	6	0.6	227	2	H75191	fibrillarlin-like p
702	6	0.6	227	2	E84403	hypothetical prote
703	6	0.6	227	2	D83543	probable transcrip
704	6	0.6	227	2	H69845	hypothetical prote
705	6	0.6	228	2	S76876	hypothetical prote
706	6	0.6	228	2	S34646	fibrillarlin-like p
707	6	0.6	228	2	T23838	hypothetical prote
708	6	0.6	228	2	AD3326	chloramphenicol ac
709	6	0.6	229	2	H64367	dolichyl-phosphate
710	6	0.6	229	2	T06204	hypothetical prote
711	6	0.6	229	2	F70568	hypothetical prote
712	6	0.6	230	2	D58892	cytochrome-c oxida
713	6	0.6	230	2	A64387	fibrillarlin-like p
714	6	0.6	230	2	S34645	fibrillarlin-like p
715	6	0.6	230	2	AF3113	transcription regu
716	6	0.6	231	2	T03765	glutathione transf
717	6	0.6	231	2	A39083	Mg2+-transporting
718	6	0.6	231	2	AD0325	urease accessory p
719	6	0.6	232	2	F64363	ribosomal protein
720	6	0.6	232	2	D85654	probable urease ac
721	6	0.6	232	2	A97748	vi18 protein limp
722	6	0.6	232	2	B90230	hypothetical prote
723	6	0.6	232	2	F97801	hypothetical prote
724	6	0.6	233	2	E72738	probable prolifera
725	6	0.6	233	2	G72527	probable fibrillar
726	6	0.6	234	2	S28482	rib protein - Vib
727	6	0.6	234	2	T30427	probable apoptos
728	6	0.6	234	2	T21543	hypothetical prote
729	6	0.6	234	2	G98173	probable transcrip
730	6	0.6	235	2	T08467	aspartate racemase
731	6	0.6	235	2	T44679	spot protein homol
732	6	0.6	235	2	A98286	hypothetical prote
733	6	0.6	235	2	H70848	probable maturase
734	6	0.6	235	2	T20337	hypothetical prote
735	6	0.6	235	2	E86127	hypothetical prote
736	6	0.6	235	2	AB1255	transport proteins
737	6	0.6	236	2	T22003	hypothetical prote
738	6	0.6	236	2	I46860	MHC RLA - rabbit (
739	6	0.6	236	2	S17932	RRNA-N-glycosidase
740	6	0.6	236	2	B70453	hypothetical prote
741	6	0.6	236	2	T43237	protein-S-isopreny
742	6	0.6	236	2	C69060	hypothetical prote
743	6	0.6	236	2	C69396	hypothetical prote
744	6	0.6	236	2	AE2268	site-specific DNA-
745	6	0.6	237	2	T47012	hypothetical prote
746	6	0.6	237	2	AD2567	two-component resp
747	6	0.6	237	2	S65313	probable membrane
748	6	0.6	237	2	G70066	capsular polysacch
749	6	0.6	237	2	H81083	hypothetical prote
750	6	0.6	237	2	D87373	hypothetical prote
751	6	0.6	237	2	T36804	probable transcrip
752	6	0.6	237	2	T24407	hypothetical prote
753	6	0.6	237	2	AE0237	ABC transporter, A
754	6	0.6	238	2	S41313	hypothetical prote
755	6	0.6	238	2	H81859	probable periplasm
756	6	0.6	238	2	AE2027	hypothetical prote
757	6	0.6	238	2	C81145	transcription regu
758	6	0.6	239	2	T45806	hypothetical prote
759	6	0.6	239	2	G81369	probable membrane
760	6	0.6	239	2	A75585	transcription regu
761	6	0.6	239	2	B95374	hypothetical prote
762	6	0.6	240	1	Z12M92	19k zein precursor
763	6	0.6	240	2	B99667	probable oxidoredu
764	6	0.6	240	2	F85517	probable oxidoredu
765	6	0.6	240	2	D95420	probable oxidoredu
766	6	0.6	240	2	A22831	19k zein precursor
767	6	0.6	240	2	H71435	hypothetical prote
768	6	0.6	241	2	C90532	ribonuclease III (
769	6	0.6	241	2	F71478	probable pseudouri
770	6	0.6	241	2	E81741	ribosomal large ch
771	6	0.6	241	2	E64386	conserved hypothet
772	6	0.6	241	2	A75065	hypothetical prote
773	6	0.6	241	2	D71167	hypothetical prote
774	6	0.6	241	2	A86651	glycosyltransferas
775	6	0.6	242	2	G82206	hypothetical prote
776	6	0.6	242	2	E88250	protein T21B10.5 (
777	6	0.6	242	2	F84315	cobalt transport A
778	6	0.6	243	2	S46732	hypothetical prote
779	6	0.6	243	2	S56405	hypothetical 26.6k
780	6	0.6	243	2	D91273	hypothetical prote
781	6	0.6	243	2	D86114	hypothetical prote
782	6	0.6	243	2	A11049	probable tRNA/rRNA
783	6	0.6	243	2	T12485	hypothetical prote
784	6	0.6	243	2	B95885	hypothetical prote
785	6	0.6	244	2	S50685	hypothetical prote
786	6	0.6	244	2	F95262	hypothetical prote
787	6	0.6	244	2	AB3524	transcription regu
788	6	0.6	245	2	S57550	hypothetical prote
789	6	0.6	245	2	F71887	hypothetical prote
790	6	0.6	245	2	A72275	hypothetical prote
791	6	0.6	245	2	H82202	hypothetical prote
792	6	0.6	246	2	G82819	acetoacetyl-CoA re
793	6	0.6	246	2	F70062	hypothetical prote
794	6	0.6	246	2	E64627	hypothetical prote
795	6	0.6	246	2	E69230	hypothetical prote
796	6	0.6	246	2	T12585	DC3 promoter-bind
797	6	0.6	246	2	B90224	hypothetical prote
798	6	0.6	246	2	A43579	Vmp7 protein homol
799	6	0.6	247	2	A99794	urase-associated
800	6	0.6	247	2	S67685	hypothetical prote
801	6	0.6	247	2	H82348	hypothetical prote
802	6	0.6	247	2	H86844	glutamine ABC tran
803	6	0.6	247	2	T14818	leucine-rich repaa
804	6	0.6	248	1	S52234	3',5'-cyclic-nucle
805	6	0.6	248	2	AC3010	gramicidin S biosy
806	6	0.6	248	2	G83478	probable permease
807	6	0.6	248	2	A98143	response regulator
808	6	0.6	248	2	AB3145	two component resp
809	6	0.6	248	2	C97439	sucrose-phosphatase
810	6	0.6	248	2	AF2657	hydrolyase (limpor
811	6	0.6	249	2	T21169	hypothetical prote
812	6	0.6	249	2	T52142	RING finger protei
813	6	0.6	249	2	AB3799	hypothetical prote
814	6	0.6	249	2	D85693	unknown protein en
815	6	0.6	250	2	T25582	hypothetical prote
816	6	0.6	250	2	F70153	conserved hypothet
817	6	0.6	250	2	B91020	hypothetical prote
818	6	0.6	250	2	D85864	hypothetical prote
819	6	0.6	250	2	AC2026	hypothetical prote
820	6	0.6	250	2	F86761	nucleotidase [limp
821	6	0.6	251	2	C83053	triisephosphate 1s
822	6	0.6	251	2	AB3390	arginine/ornithine
823	6	0.6	251	2	T11549	MAP binding protei
824	6	0.6	251	2	T14456	AAR box protein h
825	6	0.6	251	2	S73162	hypothetical prote
826	6	0.6	251	2	G97579	ABC transporter AT
827	6	0.6	251	2	AE2800	hypothetical prote
828	6	0.6	251	2	H87692	conserved hypothet
829	6	0.6	251	2	G75345	serine/threonine-p
830	6	0.6	251	2	A10100	probable carbonic
831	6	0.6	251	2	AH2181	hypothetical prote
832	6	0.6	252	2	F83098	probable short-cha

833	6	0.6	252	2	146859	MHC RLA - rabb1t (906	6	0.6	266	2	B39260	MHC class II histo
834	6	0.6	252	2	F39534	floral homeotic pr	907	6	0.6	266	2	E64955	fly protein precu
835	6	0.6	252	2	A95071	amino acid ABC tra	908	6	0.6	266	2	C90961	probable periplasm
836	6	0.6	252	2	A34877	C4b-binding protel	909	6	0.6	266	2	C85809	probable periplasm
837	6	0.6	252	2	JC2399	pms4 homolog misma	910	6	0.6	266	2	A46552	chlorophyll a/b-bl
838	6	0.6	252	2	AH1482	hypothetical prote	911	6	0.6	266	2	H69189	hypothetical prote
839	6	0.6	253	2	TS0942	D1b protein (limpo	912	6	0.6	266	2	E71612	ribosomal protein
840	6	0.6	253	2	JC5014	2,4-dichlorophenol	913	6	0.6	267	2	C64995	hypothetical prote
841	6	0.6	253	2	SA9055	thioesterase - Str	914	6	0.6	267	2	AH0792	probable 2,4-dihyd
842	6	0.6	253	2	S28541	RNA N-glycosidase	915	6	0.6	267	2	I40327	bat protein - Bord
843	6	0.6	253	2	S28539	RNA N-glycosidase	916	6	0.6	268	2	F86680	prophage pil prote
844	6	0.6	253	2	S28542	RNA N-glycosidase	917	6	0.6	268	2	S45091	hypothetical prote
845	6	0.6	253	2	S29931	RNA N-glycosidase	918	6	0.6	268	2	E75607	2-oxo-hepta-3-ene-
846	6	0.6	253	2	C36789	hypothetical prote	919	6	0.6	268	2	A64361	hypothetical prote
847	6	0.6	253	2	A90558	hypothetical prote	920	6	0.6	269	2	DS0171	chlorophyll a/b-bl
848	6	0.6	254	2	C98274	nikfz protein (Auj	921	6	0.6	269	2	D90141	hypothetical prote
849	6	0.6	254	2	F90107	60S ribosomal prot	922	6	0.6	269	2	I39961	hypothetical prote
850	6	0.6	254	2	A64437	hypothetical prote	923	6	0.6	270	2	D70171	ABC transporter, A
851	6	0.6	255	2	T12074	urate oxidase (EC	924	6	0.6	270	2	AH2455	hypothetical prote
852	6	0.6	255	2	AB3500	phosphate regulon	925	6	0.6	271	2	G89809	hypothetical prote
853	6	0.6	255	2	T18206	mxn protein - Bac	926	6	0.6	271	2	D83059	phosphatidylserine
854	6	0.6	255	2	T03299	expansin 3 - rice	927	6	0.6	271	2	AC3527	aldenhyde dehydrog
855	6	0.6	255	2	C75527	conserved hypotet	928	6	0.6	272	2	S77576	oligopeptide trans
856	6	0.6	255	2	AE3551	tetratricopeptide	929	6	0.6	272	2	D86482	protein F5j5.3 [lm
857	6	0.6	256	2	D97086	formate/nitrite fa	930	6	0.6	272	2	T25044	hypothetical prote
858	6	0.6	256	2	AB0348	hypothetical prote	931	6	0.6	272	2	AC1916	hypothetical prote
859	6	0.6	257	2	A10935	acetylglutamate kl	932	6	0.6	273	1	H69744	conserved hypotet
860	6	0.6	257	2	T64159	hypothetical prote	933	6	0.6	273	2	E72231	ATP synthase Fl, S
861	6	0.6	257	2	B90194	phosphate ABC tran	934	6	0.6	273	2	T48810	pall related prote
862	6	0.6	257	2	S33314	beta-etherase ligf	935	6	0.6	273	2	T12793	hypothetical prote
863	6	0.6	257	2	D70341	hypothetical prote	936	6	0.6	273	2	B86788	LYSR family transc
864	6	0.6	258	1	VHVW1	nucleoprotein N -	937	6	0.6	273	2	T19174	hypothetical prote
865	6	0.6	258	1	VHVW1	nucleoprotein N -	938	6	0.6	273	2	AB0848	Iron transport pro
866	6	0.6	258	1	JO1876	nucleoprotein N -	939	6	0.6	274	2	AB3583	hypothetical prote
867	6	0.6	258	1	S39747	ywfN protein - Bac	940	6	0.6	274	2	H88690	protein F4IH10.8 [
868	6	0.6	258	1	T36303	probable indolegly	941	6	0.6	274	2	F83601	hypothetical prote
869	6	0.6	258	2	S17105	nucleoprotein N -	942	6	0.6	274	2	B70901	hypothetical prote
870	6	0.6	258	2	SA7242	gene N protein - t	943	6	0.6	275	2	A87487	phosphatidate cyt
871	6	0.6	258	2	A81950	hypothetical prote	944	6	0.6	275	2	AD3383	hypothetical prote
872	6	0.6	259	2	T44461	transcription init	945	6	0.6	275	2	T43119	hypothetical prote
873	6	0.6	259	2	S32432	toxin cytB - Bacil	946	6	0.6	275	2	E93231	pur operon repress
874	6	0.6	260	1	CRMS2	carbonate dehydrat	947	6	0.6	275	2	G98095	activator of purin
875	6	0.6	260	2	H81297	Fdh protein Cj150	948	6	0.6	276	2	T09131	chitinase (EC 3.2.
876	6	0.6	261	2	D71508	probable glutamine	949	6	0.6	276	2	G69023	nucleotide-binding
877	6	0.6	261	2	T16650	hypothetical prote	950	6	0.6	276	2	AB4361	hypothetical prote
878	6	0.6	262	1	SA8029	enoyl-lacyl-carrie	951	6	0.6	276	2	C88650	protein C09G12.3 [
879	6	0.6	262	1	B43729	enoyl-lacyl-carrie	952	6	0.6	276	2	D83786	glycerol uptake fa
880	6	0.6	262	2	E90861	enoyl-lacyl-carrie	953	6	0.6	277	2	S34146	porin porI, plasti
881	6	0.6	262	2	H85757	enoyl-lacyl-carrie	954	6	0.6	277	2	S75973	hypothetical prote
882	6	0.6	262	2	AD0656	enoyl-lacyl-carrie	955	6	0.6	277	2	C90077	conserved hypotet
883	6	0.6	262	2	C97651	psbB protein (limp	956	6	0.6	278	2	C43670	integral membrane
884	6	0.6	262	2	C75051	hydrolyase related	957	6	0.6	278	2	S77601	hypothetical prote
885	6	0.6	262	2	G98343	hypothetical prote	958	6	0.6	278	2	T08252	probable transpos
886	6	0.6	262	2	AH2938	hypothetical prote	959	6	0.6	279	1	S0M7V	thermistase (EC 3.4
887	6	0.6	262	2	A65220	phosphonates trans	960	6	0.6	279	2	G95195	sugar ABC transpor
888	6	0.6	262	2	H91264	ATP-binding compon	961	6	0.6	279	2	D98062	hypothetical prote
889	6	0.6	262	2	B97070	precorrin-6x reduc	962	6	0.6	279	2	T37148	probable DNA-bind
890	6	0.6	262	2	E86105	ATP-binding compon	963	6	0.6	280	2	AB2875	exopolysaccharide
891	6	0.6	263	2	C81667	amino acid ABC tra	964	6	0.6	280	2	E71288	probable hydrogena
892	6	0.6	263	2	E97774	hypothetical prote	965	6	0.6	280	2	C91175	hypothetical prote
893	6	0.6	263	2	T16399	hypothetical prote	966	6	0.6	280	2	C86021	hypothetical prote
894	6	0.6	263	2	S73719	MG313 homolog H08	967	6	0.6	280	2	A48488	saccharide biosynt
895	6	0.6	263	2	AE0140	molybdenum transp	968	6	0.6	281	2	T09653	delta-1-pyrroline-
896	6	0.6	264	2	T31271	2-hydroxy-pent-2,4	969	6	0.6	281	2	A13196	transcription regu
897	6	0.6	264	2	A71367	probable holoctoc	970	6	0.6	281	2	S34545	hypothetical prote
898	6	0.6	264	2	F72577	hypothetical prote	971	6	0.6	281	2	E84637	hypothetical prote
899	6	0.6	265	2	T13116	protein gp30 - pha	972	6	0.6	281	2	S63225	ribosomal protein
900	6	0.6	265	2	H81155	conserved hypotet	973	6	0.6	281	2	F82670	general secretory
901	6	0.6	265	2	D90320	hypothetical prote	974	6	0.6	281	2	AE3523	icc protein [limpor
902	6	0.6	265	2	D83387	hypothetical prote	975	6	0.6	282	2	C90016	hypothetical prote
903	6	0.6	265	2	AS9141	silaflin slp1 pre	976	6	0.6	282	2	E82352	conserved hypotet
904	6	0.6	266	1	F71701	hypothetical prote	977	6	0.6	282	2	F69880	conserved hypotet
905	6	0.6	266	2	A39260	MHC class II histo	978	6	0.6	282	2	S47968	casein kinase II (

979 6 0.6 282 2 H83651 hypothetical prote
980 6 0.6 282 2 B71601 probable integral
981 6 0.6 282 2 T50833 nematode resistanc
982 6 0.6 283 2 S05205 rRNA N-glycosidase
983 6 0.6 283 2 G70376 conserved hypothet
984 6 0.6 283 2 T27423 hypothetical prote
985 6 0.6 283 2 F91285 hypothetical prote
986 6 0.6 283 2 E4348 succinoglycan bios
987 6 0.6 284 2 AH3570 nickel transport s
988 6 0.6 284 2 B64694 conserved hypothet
989 6 0.6 284 2 T39813 conserved hypothet
990 6 0.6 284 2 E95008 conserved hypothet
991 6 0.6 284 2 B97880 conserved hypothet
992 6 0.6 284 2 A70976 hypothetical prote
993 6 0.6 284 2 E82868 hypothetical prote
994 6 0.6 284 2 AD0987 probable phosphosu
995 6 0.6 285 2 C97279 thiorodoxin reduct
996 6 0.6 285 2 I50995 connexin 32.2 - At
997 6 0.6 285 2 G83233 conserved hypothet
998 6 0.6 285 2 T21611 hypothetical prote
999 6 0.6 286 2 AE1200 3-hydroxyisobutyra
1000 6 0.6 286 2 D95401 probable oxidoredu

ALIGNMENTS

RESULT 1

T27880 hypothetical protein ZK520.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jun-2000
C:Accession: T27880

R:Steward, C.

submitted to the EMBL Data Library, March 1997

A:Reference number: Z20434

A:Accession: T27880

A:Status: Preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-519 <MIL>

A:Cross-references: EMBL:Z92822; PIDN:CAB07299.1; GSPDB:GN00021; CESP:ZK520.1

A:Experimental source: clone ZK520

C:Genetics:

A:Gene: CESP:ZK520.1

A:Map position: 3

A:Introns: 12/1; 69/3; 106/2; 178/3; 240/3

C:Superfamily: Caenorhabditis elegans hypothetical protein ZK520.1

Query Match 0.8%; Score 9; DB 2; Length 519;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1009 SAIOFLVMS 1017
DB 293 SAIOFLVMS 301

RESULT 2

T17821

hypothetical protein A322L - Chlorella virus PBCV-1

C:Species: Chlorella virus PBCV-1

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 04-Mar-2000

C:Accession: T17821

R:Graves, M.V.; Van Etten, J.L.

submitted to the EMBL Data Library, May 1999

A:Reference number: Z18806

A:Accession: T17821

A:Status: Preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-176 <GRA>

A:Cross-references: EMBL:U42580; NID:g4028896; PIDN:AAC96690.1

A:Experimental source: Specific host Chlorella strain NC64A

C:Genetics:
A:Note: A322L
C:Superfamily: Chlorella virus PBCV-1 hypothetical protein A322L

Query Match 0.7%; Score 8; DB 2; Length 176;
Best Local Similarity 100.0%; Pred. No. 5.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 628 KDTIOGAK 635
DB 103 KDTIOGAK 110

RESULT 3

A84283 hypothetical protein Vng1276c [imported] - Halobacterium sp. NRC-1

C:Species: Halobacterium sp. NRC-1

C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001

C:Accession: A84283

R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky

R:Lechhauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Ja

Jung, K.H.; Alam, M.; Freitas, T.

Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000

A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ehardt, H.; Lowe, T.M.;

A:Title: Genome sequence of Halobacterium species NRC-1.

A:Reference number: A84160; MUID:20504483

A:Accession: A84283

A:Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-241 <STO>

A:Cross-references: GB:AE004437; NID:g10580799; PIDN:AAG19629.1; GSPDB:GN00138

C:Genetics:

A:Gene: VNG1276C

Query Match 0.7%; Score 8; DB 2; Length 241;
Best Local Similarity 100.0%; Pred. No. 7.3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 788 AALEMRD 795
DB 27 AALEMRD 34

RESULT 4

A83648

flagellar motor switch protein flm [imported] - Brucella melitensis (strain 16M)

C:Species: Brucella melitensis

C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002

C:Accession: A83648

R:DelVecchio, V.G.; Kapral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanov

R:DelVecchio, V.G.; Kapral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanov

Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002

A:Title: The genome sequence of the facultative intracellular pathogen Brucella melit

A:Reference number: AD3252; PMID:1175688

A:Accession: A83648

A:Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-316 <KUR>

A:Cross-references: GB:AE008918; PIDN:AAL54352.1; PID:g17985334; GSPDB:GN00191

A:Experimental source: strain 16M

C:Genetics:

A:Gene: BME11110

Query Match 0.7%; Score 8; DB 2; Length 316;
Best Local Similarity 100.0%; Pred. No. 9.3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 760 NLAGHIA 767
|||||||

Db 135 NLAGH14 142

RESULT 5

KIBSGM

phosphoglycerate kinase (EC 2.7.2.3) - *Bacillus megaterium*C:Species: *Bacillus megaterium*

C>Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 21-Jan-2000

C:Accession: S13125; J01954

R:Schlaepfer, B.S.; Brniant, G.; Brniant, G.; Zuber, H.

Nucleic Acids Res. 18, 6423, 1990

A:Title: Nucleotide sequence of the phosphoglycerate kinase gene from *Bacillus megaterium*

A:Reference number: S13125; M01D:91057129

A:Accession: S13125

A:Molecule type: DNA

A:Residues: 1-394 <SCH>

A:Cross-references: EMBL:X54519; NID:g39642; PIDN:CAA38375.1; PID:g39643

A:Experimental source: DSM 319

R:Schlaepfer, B.S.; Zuber, H.

Gene 122, 53-62, 1992

A:Title: Cloning and sequencing of the genes encoding glyceraldehyde-3-phosphate dehydrogenase and corresponding sequences from thermophilic *Bacillus stearothermophilus*.

A:Reference number: J01952; M01D:93083995

A:Accession: J01954

A:Molecule type: DNA

A:Residues: 1-394 <SC2>

A:Cross-references: GB:M87647; NID:g143315; PIDN:AA73203.1; PID:g143318

A:Experimental source: strain DSM319

C:Genetics:

A:Gene: pgk

C:Superfamily: phosphoglycerate kinase

C:Keywords: ATP; gluconeogenesis; glycolysis; phosphotransferase

F:201,323/Binding site: ATP (Lys, Glu) #status predicted

Query Match

Best Local Similarity 0.7%; Score 8; DB 1; Length 394;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 632 QGAKV14 639

Db 50 QGAKV14 57

RESULT 6

AC0575

hypothetical membrane protein p43 STY0637 [imported] - *Salmonella enterica* subsp. *enterica*C:Species: *Salmonella enterica* subsp. *enterica* serovar TyphA:Note: this species has also been called *Salmonella typhi*

C>Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 27-Nov-2001

C:Accession: AC0575

R:Perkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Moule, S.; O'Garra, P.

Nature 413, 848-852, 2001

A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; A:Title: Complete genome sequence of a multiple drug resistant *Salmonella enterica* serovar

A:Reference number: AB0502; PMID:11677608

A:Accession: AC0575

A:Molecule type: DNA

A:Residues: 1-414 <PAR>

A:Cross-references: GB:AL513382; PIDN:CAD05069.1; PID:g16501844; GSPDB:GN00176

C:Genetics:

A:Gene: STY0637

C:Superfamily: Streptomyces lividans chloramphenicol resistance protein

Query Match

Best Local Similarity 0.7%; Score 8; DB 2; Length 414;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 148 VGSFLAVG 155

Db 296 VGSFLAVG 303

RESULT 7

E87233

probable secreted protein [imported] - *Mycobacterium leprae*C:Species: *Mycobacterium leprae*

C>Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001

C:Accession: E87233

R:Coile, S.T.; Eljmeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; R.; Davies, R.M.; Devlin, K.; Duthoy, S.; Fellwell, T.; Fraser, A.; Hamlin, N.; Holto

eam, M.A.; Rutherford, K.M.

Nature 409, 1007-1011, 2001

A:Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; A:Title: Massive gene decay in the leprosy bacillus.

A:Reference number: A86909; M01D:21128732; PMID:11234002

A:Accession: E87233

A:Molecule type: DNA

A:Residues: 1-519 <STO>

A:Cross-references: GB:AL450380; NID:g13093825; PIDN:CAC32123.1; GSPDB:GN00147

C:Genetics:

A:Gene: mceIC

Query Match

Best Local Similarity 0.7%; Score 8; DB 2; Length 519;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 561 FPDNGTR 568

Db 346 FPDNGTR 353

RESULT 8

T29967

hypothetical protein F36A4.15 - *Caenorhabditis elegans*C:Species: *Caenorhabditis elegans*

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 02-Jun-2000

C:Accession: T29967

R:Miller, N.; Bradshaw, H.

submitted to the EMBL Data Library, April 1996

A:Description: The sequence of C. elegans cosmid F36A4.

A:Reference number: Z20713

A:Accession: T29967

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-642 <ML>

A:Cross-references: EMBL:U53333; PIDN:AA96166.1; GSPDB:GN00022; CESP:F36A4.15

A:Experimental source: strain Bristol N2; clone F36A4

C:Genetics:

A:Gene: CESP:F36A4.15

A:Map position: 4

A:Insertions: 14/3; 116/2; 146/3; 181/2; 388/3; 466/2

C:Superfamily: xanthine dehydrogenase; ferredoxin [2Fe-2S] homology

Query Match

Best Local Similarity 0.7%; Score 8; DB 2; Length 642;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1057 FEGEKRYL 1064

Db 282 FEGEKRYL 289

RESULT 9

T34618

NADH dehydrogenase (ubiquinone) (EC 1.6.5.3) chain nuoG - *Streptomyces coelicolor*C:Species: *Streptomyces coelicolor*

C>Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 04-Mar-2000

C:Accession: T34618

R.Saunders, D.C.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.
submitted to the EMBL Data Library, June 1999
A:Reference number: 221549
A:Accession: T34618
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-843 <SAU>
A:Cross-references: EMBL:AL078618; PIDN:CAB44525.1; GSPDB:GN00070; SCOEDB:SC10A7.15C
A:Experimental source: strain A3(2)
C:Genetics:
A:Gene: SCOEDB:SC10A7.15C
C:Superfamily: NADH dehydrogenase (ubiquinone) chain 11
C:Keywords: NAD; oxidoreductase

Query Match 0.7%; Score 8; DB 2; Length 843;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 104 KGDVLA 111
|||||||
DB 775 KGDVLA 782

RESULT 10
T47659
splicosomal-like protein - Arabidopsis thaliana
N:Alternat names: protein T26112.100; protein T26112.80
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 19-May-2000
R:Monfort, A.; Casacuberta, E.; Puigdomenech, P.; Mewes, H.W.; Lemcke, K.; Mayer, K.F.X.
submitted to the Protein Sequence Database, February 2000
A:Reference number: Z24471
A:Accession: T47659
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1214 <MON>
A:Cross-references: EMBL:AL132954
A:Experimental source: cultivar Columbia; BAC clone T26112
A:Accession: T47661
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1214 <MO2>
A:Cross-references: EMBL:AL132954
A:Experimental source: cultivar Columbia; BAC clone T26112
C:Genetics:
A:Map position: 3
A:Introns: 1005/3; 1169/3
A>Note: T26112.80; T26112.100

Query Match 0.7%; Score 8; DB 2; Length 1214;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 152 LAVGVKG 159
|||||||
DB 902 LAVGVKG 909

RESULT 11
T32625
hypothetical protein F15E6.6 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 18-Feb-2000
C:Accession: T32625
R:Miller, N.; Stelliyes, L.; Bradshaw, H.; Keppler, D.
submitted to the EMBL Data Library, December 1997
A:Description: The sequence of C. elegans cosmid F15E6.
A:Reference number: Z21202
A:Accession: T32625
A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA
A:Residues: 1-1238 <MI>
A:Cross-references: EMBL:AF088614; PIDN:AB92058.1; GSPDB:GN00022; CESP:F15E6.6
A:Experimental source: strain Bristol NZ; clone F15E6
C:Genetics:
A:Gene: CESP:F15E6.6
A:Map position: 4
A:Introns: 34/1; 109/3; 225/3; 271/1; 312/1; 345/3; 412/3; 517/3; 551/1; 616/3; 709/3
C:Superfamily: xanthine dehydrogenase; ferredoxin [2Fe-2S] homology
C:Keywords: 2Fe-2S; metalloprotein
F:43,48,51,73/Binding site: 2Fe-2S cluster (Cys) (covalent) #status predicted

Query Match 0.7%; Score 8; DB 2; Length 1238;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1057 FEGERKYL 1064
|||||||
DB 878 FEGERKYL 885

RESULT 12
G98275
hypothetical protein AGR_L_2329 [imported] - Agrobacterium tumefaciens (strain C58, C
C:Species: Agrobacterium tumefaciens
C>Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 11-Jan-2002
C:Accession: G98275
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quirillo, B.; Gold
A.; Liu, F.; Moliam, C.; Allinger, M.; Doughy, D.; Scott, C.; Lappas, C.; Matkeitz,
Science 294, 2353-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium
A:Reference number: A97359; PMID:11743194
A:Accession: G98275
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1520 <KUR>
A:Cross-references: GB:AE007870; PIDN:AAK89729.1; PID:q15159645; GSPDB:GN00170
C:Genetics:
A:Gene: AGR_L_2329
A:Map position: linear chromosome

Query Match 0.7%; Score 8; DB 2; Length 1520;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 633 GAKVILAG 640
|||||||
DB 1175 GAKVILAG 1182

RESULT 13
AF3008
polyketide synthase Atu3672 [imported] - Agrobacterium tumefaciens (strain C58, Dupon
C:Species: Agrobacterium tumefaciens
C>Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 11-Jan-2002
C:Accession: AF3008
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kuyavlin, T.; Levy, R.; Li, M.; McCl
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kam
ster, E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; PMID:11743193
A:Accession: AF3008
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1520 <KUR>
A:Cross-references: GB:AE008689; PIDN:AA144484.1; PID:q17742089; GSPDB:GN00187
A:Experimental source: strain C58 (Dupont)
C:Genetics:

A:Gene: Atu3672
A:Map position: linear chromosome

Query Match 0.7%; Score 8; DB 2; Length 1520;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 633 GAKVLAG 640
|||||
DB 1175 GAKVLAG 1182

RESULT 14

A96026
probable transposase of insertion sequence ISKm2011-2, orfB N-terminus protein [imported]
C:Species: Sinorhizobium meliloti
C:Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001
C:Accession: A96026
R:Finan, T.M.; Weidner, S.; Wong, K.; Buhrmaster, J.; Chain, P.; Vorholter, F.J.; Hernan
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
A:Title: The complete sequence of the 1,683-kb pSymb megaplasmid from the N2-fixing endo
A:Reference number: A93842; MUID:2136508; PMID:11481431
A:Accession: A96026
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-39 <KUR>
A:Cross-references: GB:AL591985; PIDN:CAC49874.1; PID:915141362; GSPDB:GN00167
A:Experimental source: strain 1021, megaplasmid pSymb
R:Galibert, F.; Finan, T.M.; Long, S.R.; Publer, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
hebut, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yen, K.
A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A:Reference number: A96039; MUID:21368234; PMID:11474104
A:Contents: annotation
C:Genetics:
A:Gene: Trm2011-2b-1; Smb21702
A:Genome: plasmid

Query Match 0.7%; Score 7; DB 2; Length 39;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 568 RLVFIDE 574
|||||
DB 32 RLVFIDE 38

RESULT 15

S14968
hypothetical NADH dehydrogenase chain 4 homolog - Synechocystis sp. (strain PCC 6803)
N:Alternate names: ndhd protein-like protein
C:Species: Synechocystis sp.
A:Variety: PCC 6803
C:Date: 15-Oct-1998 #sequence_revision 15-Oct-1998 #text_change 02-Jun-2000
C:Accession: S14968
R:Anderson, S.L.; McIntosh, L.
Plant Mol. Biol. 16, 487-499, 1991
A:Title: Partial conservation of the 5' ndhE-psac-ndhd 3' gene arrangement of chloroplas
oplasts.
A:Reference number: S14966; MUID:91329685
A:Accession: S14968
A:Molecule type: DNA
A:Residues: 1-91 <AND>
A:Cross-references: EMBL:X53842; NID:947594; PIDN:CA37837.1; PID:9581748
A:Experimental source: PCC 6803
A:Note: the authors translated the initiation codon GTG as Val
C:Comment: This is the hypothetical translation of a sequence that was not reported as a
co chloroplast (see PIR:DEMTN4).

C:Genetics:
A:Start codon: GTG

Query Match 0.7%; Score 7; DB 4; Length 91;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 230 KMDDRTS 236
|||||
DB 62 KMDDRTS 68

Search completed: May 21, 2002, 14:55:07
Job time: 186 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 21, 2002, 14:52:21 ; Search time 15.19 Seconds
(without alignments)
2737.643 Million cell updates/sec

Title: US-09-729-653-2_COPY_1_1074

Perfect score: 1074

Sequence: 1 HSLIGRCSRLGIDGNVAC.....LYFEGEKRYLQAGREFLLCG 1074

Scoring table: OLIGO

Searched: 105224 seqs, 38719550 residues

Word size : 0

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: listing first 1000 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	ID	Description
1	8	0.7	394 1	PGK_BACME
2	8	0.7	843 1	NUOG_STRCO
3	7	0.7	15 1	LCK_DROME
4	7	0.7	120 1	YACL_ECOLI
5	7	0.7	131 1	RT17_YEAST
6	7	0.7	144 1	RL15_ECOLI
7	7	0.7	144 1	RL15_HAEIN
8	7	0.7	157 1	FMO_MORBO
9	7	0.7	159 1	FMI_MORBO
10	7	0.7	159 1	MLE_TODPA
11	7	0.7	173 1	POPS_YEAST
12	7	0.7	182 1	UCRT_RHORI
13	7	0.7	190 1	YEIP_ECOLI
14	7	0.7	196 1	SODF_TETPY
15	7	0.7	217 1	DFRA_MEDSA
16	7	0.7	219 1	R1B7_ARCFU
17	7	0.7	225 1	CSMJ_CHUTE
18	7	0.7	249 1	BIAB_BACFR
19	7	0.7	262 1	MURI_BUCAT
20	7	0.7	267 1	YG95_HAEIN
21	7	0.7	287 1	V033_FOWPV
22	7	0.7	295 1	RFBA_XANCP
23	7	0.7	302 1	MEUR_ECOLI
24	7	0.7	331 1	LEUI_THETH
25	7	0.7	332 1	TWPI_YEAST
26	7	0.7	336 1	G3P_ASPNG
27	7	0.7	337 1	G3P_CRYPA
28	7	0.7	338 1	G3P_NEUCR
29	7	0.7	351 1	MTD5_DACSA
30	7	0.7	359 1	VALI_TYLCM
31	7	0.7	359 1	VALI_TYLCU
32	7	0.7	364 1	VMAT_NOVA
33	7	0.7	364 1	VMAT_NDVB

34	7	0.7	370 1	VE2_HPV39	P24630 human papil
35	7	0.7	386 1	RIR2_BRARE	P79733 brachydanio
36	7	0.7	394 1	PGK_BACST	P18912 bacillus st
37	7	0.7	396 1	DXR_DEIRA	O9RU84 deinoceccus
38	7	0.7	401 1	PGKT_WHEAT	P12783 triticum ae
39	7	0.7	402 1	DFP_STNY3	P73881 synechocyst
40	7	0.7	443 1	SGN2_HUMAN	Q15647 homo sapien
41	7	0.7	448 1	TX19_HUMAN	O60806 homo sapien
42	7	0.7	450 1	SR54_MYCPN	P75054 mycoplasma
43	7	0.7	485 1	GLME_CLOTT	O05509 clostridium
44	7	0.7	535 1	D2_DICDI	P18142 dicystostell
45	7	0.7	536 1	PGCK_RHIME	P43085 rhizobium m
46	7	0.7	547 1	GUNI_BUTFI	P20847 butyryllybri
47	7	0.7	566 1	ROCB_BACSU	P39635 bacillus su
48	7	0.7	574 1	DPO1_AOUAE	O67779 aquifex aeo
49	7	0.7	576 1	PEX5_PICPA	P33392 pichia past
50	7	0.7	589 1	RESE_BACSU	P35164 bacillus su
51	7	0.7	612 1	GIDA_MYCGE	P47619 mycoplasma
52	7	0.7	629 1	HTPG_RHIME	P58477 rhizobium m
53	7	0.7	632 1	DXS_CHLMU	O9PK62 chlamydia m
54	7	0.7	640 1	DXS_CHLTR	O84335 chlamydia t
55	7	0.7	654 1	PGKT_THEMA	P36304 thermotoga
56	7	0.7	695 1	EBG1_TREPA	O83748 treponema p
57	7	0.7	702 1	SPB1_BRAJU	O82475 brassica ju
58	7	0.7	745 1	AXN_DROME	O9V407 drosophila
59	7	0.7	816 1	YM35_YEAST	O03823 saccharomyc
60	7	0.7	889 1	COPE_YEAST	P41811 saccharomyc
61	7	0.7	906 1	HELY_MYCTU	Q10701 mycobacteri
62	7	0.7	957 1	NIRB_KLEPN	O06458 klebsiella
63	7	0.7	988 1	DPOL_FOWPV	P21402 foxipox vir
64	7	0.7	993 1	YIS2_YEAST	P40562 saccharomyc
65	7	0.7	1188 1	PPSA_METJA	O57862 methanococc
66	7	0.7	1296 1	BXG_CLOBO	O60393 clostridium
67	7	0.7	1403 1	PRO_DROME	P29617 drosophila
68	7	0.7	1612 1	TP2B_CRILO	O64399 cricetus
69	7	0.7	1612 1	TP2B_MOUSE	O64511 homo musculu
70	7	0.7	1626 1	TP2B_HUMAN	O02880 homo sapien
71	7	0.7	1643 1	OMPR_RICPR	O53020 r outer mem
72	7	0.7	2090 1	N214_HUMAN	P35658 homo sapien
73	7	0.7	3305 1	APLP_MANSE	O25490 manduca sex
74	6	0.6	15 1	LEC3_AXIPO	P28888 axinella po
75	6	0.6	54 1	BAF_BORBR	O08005 bordetella
76	6	0.6	56 1	YPMA_BACSU	P54395 bacillus su
77	6	0.6	63 1	MT1_COLLI	P15786 columba liv
78	6	0.6	66 1	PSAE_FREDI	P23697 fremyella d
79	6	0.6	67 1	CSRA_MECIU	O30875 micrococcu
80	6	0.6	69 1	PSAE_STNP2	P31369 synechococc
81	6	0.6	70 1	PSAE_CYAPA	O48114 cyanophora
82	6	0.6	85 1	DYLI_SCHPO	O9U005 schizosacch
83	6	0.6	87 1	F115_MOUSE	O61075 mus musculu
84	6	0.6	87 1	YD63_ARCFU	O28908 archaeoglob
85	6	0.6	88 1	ATP9_DICDI	P37315 dicystostell
86	6	0.6	88 1	PMRD_ECOLI	P37590 escherichia
87	6	0.6	99 1	Y142_UREPA	O9P102 ureaplasma
88	6	0.6	101 1	VE4_HPV41	P27553 human papil
89	6	0.6	105 1	GLN1_METBA	P54808 methanosarc
90	6	0.6	107 1	FKB1_BOVIN	P18023 bos taurus
91	6	0.6	107 1	FKB1_HUMAN	P20071 homo sapien
92	6	0.6	107 1	FKB1_MOUSE	P26883 mus musculu
93	6	0.6	107 1	FKB1_XENLA	O62658 rattus norv
94	6	0.6	107 1	FKB1_XENLA	O42123 xenopus lae
95	6	0.6	107 1	GUAN_CAVPO	P70664 cavia porce
96	6	0.6	108 1	KV3A_HUMAN	P01619 homo sapien
97	6	0.6	109 1	CLP1_HILAM	O02443 helicoverpa
98	6	0.6	109 1	GUAN_PIG	P79697 sus scrofa
99	6	0.6	111 1	YW12_STRCO	P23159 streptomyce
100	6	0.6	115 1	GUAN_HUMAN	O02747 homo sapien
101	6	0.6	116 1	SMS_CHICK	P33094 gallus gall
102	6	0.6	116 1	WN5B_ALOVU	P28114 alopias vul
103	6	0.6	116 1	WN5B_EUMSK	P28118 eumeces ski
104	6	0.6	118 1	POL_HVTC4	P05560 human immun
105	6	0.6	119 1	FOLX_ECOLI	P80449 escherichia
106	6	0.6	121 1	RS10_MYCGA	O52331 mycoplasma

107	6	0.6	121	1	Y127_METUA	05791	methanococ	180	6	0.6	184	1	ESM1_RAT	P9762	rattus nov
108	6	0.6	122	1	DBH_DEIRA	Q9189	delnoco	181	6	0.6	185	1	SOMA_KATPE	P20391	katsunon
109	6	0.6	122	1	TOA2_YEAST	P2774	saccharomyc	182	6	0.6	187	1	DEF_FRBD1	P94601	fremyella d
110	6	0.6	123	1	GLN2_METBA	P54809	methanarc	183	6	0.6	187	1	SOMA_THUAL	P34747	thinnus alb
111	6	0.6	124	1	CC26_YEAST	P14724	saccharomyc	184	6	0.6	188	1	HBA_MYCLE	P09c44	mycobacteri
112	6	0.6	124	1	CHTB_VIBCH	P01556	vibrio chol	185	6	0.6	191	1	PABA_SBRMA	P06192	serratia ma
113	6	0.6	126	1	RS13_AOUAE	O66486	aquifex aeo	186	6	0.6	191	1	RR9_CHIRE	O20029	chlamydomon
114	6	0.6	126	1	Y4SK_RHISN	P55654	rhizobium s	187	6	0.6	191	1	TFAR_ECOTI	P76135	escherichta
115	6	0.6	127	1	Y731_HAEIN	P44044	haemophilus	188	6	0.6	191	1	TFAR_ECOTI	P77163	escherichta
116	6	0.6	128	1	RNPA_MYCGE	P47703	mycoplasma	189	6	0.6	192	1	Y703_HAEIN	P44832	haemophilus
117	6	0.6	129	1	RL32_ARCFU	O28371	archaeoglob	190	6	0.6	193	1	YB71_HAEIN	P44339	haemophilus
118	6	0.6	131	1	PRDC_PROVU	P20823	proteus vul	191	6	0.6	194	1	ACPD_HAEIN	P43013	haemophilus
119	6	0.6	131	1	TAP2_ANTMA	P24609	antirrhinum	192	6	0.6	194	1	H5B_XENLA	P22845	xenopus lae
120	6	0.6	132	1	Y274_HELPY	O92mg1	helicobacte	193	6	0.6	194	1	LPCA_HAEDU	O87340	haemophilus
121	6	0.6	132	1	Y274_HELPY	P56132	helicobacte	194	6	0.6	194	1	TPA_LAMB	O37140	bacterioph
122	6	0.6	133	1	THH2_ARATH	O38879	arabidopsis	195	6	0.6	195	1	NIRQ_AZOVI	P11068	azotobacter
123	6	0.6	133	1	YGLB_BUTFI	P30541	butyrivibri	196	6	0.6	195	1	H5A_XENLA	P22844	xenopus lae
124	6	0.6	134	1	YHFA_ECOLI	P24246	escherichta	197	6	0.6	196	1	ZO26_XENLA	P18746	xenopus lae
125	6	0.6	137	1	Y877_ARCFU	O29384	archaeoglob	198	6	0.6	198	1	HBA_MYCTU	O11142	mycobacteri
126	6	0.6	139	1	PSAD_SKECO	O96800	skeletonema	199	6	0.6	199	1	TRPG_PSEPU	P00901	pseudomonas
127	6	0.6	139	1	RNPA_CHLPN	O926x2	chlamydia p	200	6	0.6	200	1	SODE_BABBO	O15905	babesia bov
128	6	0.6	141	1	HBA_COLLI	P21871	columba liv	201	6	0.6	200	1	GCH2_PHOPO	P51618	photobacter
129	6	0.6	141	1	HBA_TAPTE	P01962	tapirus ter	202	6	0.6	203	1	SOMA_PAGNA	P08591	pagrus majo
130	6	0.6	141	1	R25B_SCHPO	O74391	schizosacch	203	6	0.6	203	1	VARD_CHLNU	O9pk87	chlamydia m
131	6	0.6	141	1	RL13_HELPY	O92my6	helicobacte	204	6	0.6	204	1	HIS_LEPIN	O954h8	leptospira
132	6	0.6	141	1	RL13_HELPY	P56038	helicobacte	205	6	0.6	204	1	SOMA_ACABU	O01282	acanthopagr
133	6	0.6	141	1	SSRP_UREPA	O9p87	ureaplasma	206	6	0.6	204	1	SOMA_ACABA	P45654	acanthopagr
134	6	0.6	141	1	Y496_METUA	O5719	methanococ	207	6	0.6	204	1	SOMA_LATCA	O01283	latca calca
135	6	0.6	141	1	YKLA_BACSU	O34762	bacillus su	208	6	0.6	204	1	SOMA_PERYV	O9dev3	perca flave
136	6	0.6	142	1	MM51_MYCTU	P95212	mycobacteri	209	6	0.6	204	1	SOMA_PSECK	O919m4	pseudosclae
137	6	0.6	142	1	PK40_SPIOL	P27684	spiniactia ol	210	6	0.6	204	1	SOMA_SEBSC	P87391	sebastes sc
138	6	0.6	143	1	PTMA_MYCPN	P75145	mycoplasma	211	6	0.6	204	1	SOMA_SERBU	P09539	seriola qui
139	6	0.6	144	1	HOIC_HAEIN	P43749	haemophilus	212	6	0.6	204	1	SOMA_SPAUV	P29971	spurus aura
140	6	0.6	144	1	RL15_BUCAT	P57572	buchnera ap	213	6	0.6	204	1	SOMA_THUTH	P09113	thinnus thy
141	6	0.6	145	1	Y047_ARCFU	O30189	archaeoglob	214	6	0.6	204	1	YIGZ_ECOLI	P27862	escherichta
142	6	0.6	145	1	FLIJ_BUCAT	P57179	buchnera ap	215	6	0.6	206	1	YIGZ_HAEIN	P44842	haemophilus
143	6	0.6	146	1	HBE_CHICK	P02128	gallus gall	216	6	0.6	207	1	MADD_BACHD	O9kd91	bacillus ha
144	6	0.6	146	1	Y142_AOUAE	O66337	aquifex aeo	217	6	0.6	207	1	RISA_AOUAE	O67604	aquifex aeo
145	6	0.6	147	1	RL13_STRCO	O53874	streptomyce	218	6	0.6	209	1	BE16_MOUSE	P28657	mus musculu
146	6	0.6	147	1	YJVI_YEAST	P40898	saccharomyc	219	6	0.6	209	1	COBB_SAISC	O9xsm7	saimitri sci
147	6	0.6	147	1	YMLB_BACSU	P39152	bacillus su	220	6	0.6	209	1	COEC_SCYCA	P81130	scyllorhinu
148	6	0.6	148	1	NDK_THHAC	O9h159	thermoplasm	221	6	0.6	209	1	SHR3_YEAST	O02774	saccharomyc
149	6	0.6	151	1	NPC2_HUMAN	O15668	homo sapien	222	6	0.6	209	1	VAD2_TREPA	O83539	treponema p
150	6	0.6	151	1	Y448_MYCPN	P75129	mycoplasma	223	6	0.6	210	1	YG21_YEAST	P53251	saccharomyc
151	6	0.6	152	1	NI01_HUMAN	O92982	homo sapien	224	6	0.6	212	1	RA1I_RHIEF	O54451	rhizobium e
152	6	0.6	152	1	NI01_MOUSE	O70131	mus musculu	225	6	0.6	213	1	RL1_METTV	O52704	methanococ
153	6	0.6	152	1	NI01_RAT	P70617	rattus nov	226	6	0.6	213	1	RL1_METVA	P15824	methanococ
154	6	0.6	152	1	Y000_MYCTU	O50621	mycobacteri	227	6	0.6	213	1	RL1_METVO	O9h812	methanococ
155	6	0.6	153	1	VG17_BPPHS	P15850	bacterioph	228	6	0.6	213	1	SODM_PASHA	O95679	pasteurella
156	6	0.6	154	1	ELIS_HALFO	O01381	halitosis fu	229	6	0.6	216	1	KITH_STRCO	O50519	streptomyce
157	6	0.6	154	1	FMN_MORNO	P09829	moraxella n	230	6	0.6	216	1	PAAD_STRCO	O9kyp1	streptomyce
158	6	0.6	155	1	RLM6_PETHY	O04715	petunia hyb	231	6	0.6	216	1	YI04_YEAST	P40576	saccharomyc
159	6	0.6	157	1	PUR6_METUA	O58033	methanococ	232	6	0.6	217	1	MTN_THEMA	O9x013	thermotoga
160	6	0.6	157	1	RIP4_SAPOF	P27561	saponaria o	233	6	0.6	217	1	Y875_CHLPN	O97321	chlamydia p
161	6	0.6	159	1	ECRA_ETICO	P35645	eikenella c	234	6	0.6	218	1	DNE1_CHLNU	O32001	chlamydomon
162	6	0.6	162	1	PTSN_KLEPN	P17162	klebsiella	235	6	0.6	218	1	LEP2_SYNT3	P73157	synecocyst
163	6	0.6	162	1	RRPO_LVX	P27327	lily virus	236	6	0.6	219	1	ERG2_SCHPO	P87113	schizosacch
164	6	0.6	162	1	TPX_STRGC	P42366	streptococ	237	6	0.6	219	1	FLGA_SALTY	P40131	salmonella
165	6	0.6	162	1	TPX_STRSA	P31308	streptococ	238	6	0.6	219	1	PDXH_MYCLE	O33065	mycobacteri
166	6	0.6	164	1	BFR1_MAGMG	O50171	magnetospir	239	6	0.6	219	1	RL1_METVA	P54050	methanococ
167	6	0.6	166	1	VG17_BPPH2	O55235	synecococ	240	6	0.6	220	1	BIOD_AOUAE	O66832	aquifex aeo
168	6	0.6	166	1	SC11_YEAST	P03886	bacterioph	241	6	0.6	220	1	CFIB_PETHY	P11651	petunia hyb
169	6	0.6	167	1	YMH6_CAEEL	P15367	saccharomyc	242	6	0.6	220	1	Y064_ARCFU	O31072	archaeoglob
170	6	0.6	168	1	RR7_CHIRE	P34473	caenorhabd	243	6	0.6	221	1	GC12_STRCO	O88011	streptomyce
171	6	0.6	168	1	YMH6_CAEEL	P48627	chlamydomon	244	6	0.6	221	1	YB31_MYCPN	P75267	mycoplasma
172	6	0.6	174	1	VG17_BPPZA	P08389	bacterioph	245	6	0.6	222	1	CDTA_ACTAC	O87120	actinobacil
173	6	0.6	177	1	RM16_BRANA	P49389	brassica na	246	6	0.6	222	1	Y4BK_RHISN	P55378	rhizobium s
174	6	0.6	178	1	HPRT_BUCAT	P57291	buchnera ap	247	6	0.6	222	1	YDCL_ECOLI	P76101	escherichta
175	6	0.6	179	1	APR_PASMU	O95147	arabidopsis	248	6	0.6	223	1	CDTA_HAEDU	O02522	haemophilus
176	6	0.6	180	1	APR_PASMU	P57841	pasteurella	249	6	0.6	223	1	SAMP_HUMAN	P02743	homo sapien
177	6	0.6	180	1	APR_RAT	P36972	rattus nov	250	6	0.6	223	1	YG52_XENLA	P18727	xenopus lae
178	6	0.6	182	1	VB07_VACCV	P21003	vaccinia vi	251	6	0.6	224	1	Y945_METUA	O58355	methanococ
179	6	0.6	184	1	ESM1_MOUSE	O9qy77	mus musculu	252	6	0.6	226	1	BIOD_MYCLE	P45486	mycobacteri

253	6	0.6	226	1	THRA_TOBAC	P80969 nicotiana t	326	6	0.6	284	1	PPNK_HELPY	O25944 helicobacte
254	6	0.6	226	1	THTB_TOBAC	Q9smb8 nicotiana t	327	6	0.6	285	1	Y313_MYCPN	O53064 mycoplasma
255	6	0.6	227	1	FLPA_PYRAB	O9v215 pyrococcus	328	6	0.6	287	1	CAPB_YEAST	P13317 saccharoma
256	6	0.6	227	1	FLPA_PYRHO	O57811 pyrococcus	329	6	0.6	287	1	SUD2_ARCFU	O28098 archaeoglob
257	6	0.6	228	1	FLPA_METVO	P35553 methanococc	330	6	0.6	288	1	H01_HUMAN	P09601 homo sapien
258	6	0.6	229	1	Y544_METUA	O57964 methanococc	331	6	0.6	288	1	PPNK_PORGI	O51841 porphyromon
259	6	0.6	230	1	FLPA_METJA	O58108 methanococc	332	6	0.6	290	1	ILVE_ARCFU	O29329 archaeoglob
260	6	0.6	230	1	FLPA_METJA	P35552 methanococc	333	6	0.6	291	1	RGR_BOVIN	P47803 bos taurus
261	6	0.6	230	1	UREE_YEREN	P42869 yersinia en	334	6	0.6	291	1	RGR_MOUSE	O9z2b3 mus musculu
262	6	0.6	231	1	ATWC_SALTY	P22037 salmoneilla	335	6	0.6	292	1	RIP2_SAPOR	P27595 saponaria o
263	6	0.6	231	1	COX2_LATCH	O03648 latimeria c	336	6	0.6	294	1	FDNH_ECOLI	P24184 escherichia
264	6	0.6	231	1	UREE_YERPE	O9z2tr yersinia pe	337	6	0.6	294	1	LEP_MYCTU	O10789 mycobacteri
265	6	0.6	231	1	UREE_YERPE	P52317 yersinia ps	338	6	0.6	294	1	SUOE_HUMAN	P49688 homo sapien
266	6	0.6	233	1	FLPA_AERPE	O9y9u3 aeropyrum p	339	6	0.6	294	1	SUPP_BOVIN	P50227 bos taurus
267	6	0.6	236	1	MAM4_SCHPO	P87014 s proteinf-s	340	6	0.6	295	1	SP19_YEAST	P32573 saccharomyc
268	6	0.6	236	1	RIP3_SAPOR	P27660 saponaria o	341	6	0.6	296	1	EX53_BACSU	P54161 bacillus su
269	6	0.6	236	1	RL3_MYCGA	O52332 mycoplasma	342	6	0.6	297	1	TY5Y_PNECA	P13100 pneumocysti
270	6	0.6	236	1	YH76_AOUAE	O67652 aquifex aeo	343	6	0.6	298	1	REG5_DROME	O94913 drosophila
271	6	0.6	237	1	YH9A_YEAST	P04702 zea mays (m	344	6	0.6	299	1	RIP6_SAPOR	P20656 saponaria o
272	6	0.6	240	1	ZEAG_MAIZE	P04702 zea mays (m	345	6	0.6	301	1	URID_CANLI	P34799 canavalia l
273	6	0.6	240	1	ZEAG_MAIZE	P06677 zea mays (m	346	6	0.6	302	1	FSRL_ARATH	P52839 arabidopsis
274	6	0.6	241	1	Y693_METJA	O58104 methanococc	347	6	0.6	303	1	HUXE_ECOLI	P77335 escherichia
275	6	0.6	243	1	TRPA_CYAME	O22018 cyanidiosch	348	6	0.6	303	1	RLUC_RICPR	O9z6r7 rickettsia
276	6	0.6	243	1	YHK9_YEAST	P38777 saccharomyc	349	6	0.6	303	1	V212_FOWPV	O9j323 fowlpox vir
277	6	0.6	243	1	YJFH_ECOLI	P39290 escherichia	350	6	0.6	303	1	YJTE_ECOLI	P39376 escherichia
278	6	0.6	244	1	CY1_CHRVI	O31216 chromatium	351	6	0.6	304	1	UL49_HSEVB	P28960 equine herp
279	6	0.6	244	1	YB12_YEAST	P40098 saccharomyc	352	6	0.6	305	1	CAH5_HUMAN	P35518 homo sapien
280	6	0.6	245	1	IR6_YEAST	O12522 saccharomyc	353	6	0.6	305	1	ENDA_ARCFU	O29362 archaeoglob
281	6	0.6	249	1	PCW3_AERPE	O9y9ez aeropyrum p	354	6	0.6	305	1	MINA_THEMA	O29362 archaeoglob
282	6	0.6	251	1	ABCX_PORPU	P51441 porphyra pu	355	6	0.6	305	1	YB61_SYNY3	P73585 thermotoga
283	6	0.6	251	1	TPIS_PSEAE	O9hv51 pseudomonas	356	6	0.6	306	1	CD80_MOUSE	O00609 mus musculu
284	6	0.6	251	1	YSCI_SYRGC	P42360 streptococc	357	6	0.6	306	1	HEM3_HELPY	P56149 mus musculu
285	6	0.6	252	1	AGL6_ARATH	P29386 arabidopsis	358	6	0.6	307	1	YDJ2_SYNP7	O33701 synchococc
286	6	0.6	252	1	CABR_HUMAN	P20851 homo sapien	359	6	0.6	308	1	GLS1_BACHD	O9K181 bacillus ha
287	6	0.6	253	1	RIP5_SAPOR	O41389 saponaria o	360	6	0.6	308	1	SAL_STIAS	O9Pw8 silurus aso
288	6	0.6	253	1	RIP7_SAPOR	O41391 saponaria o	361	6	0.6	308	1	URIC_CANLI	P34798 canavalia l
289	6	0.6	253	1	VG29_HSVI1	O00151 ictaluriid h	362	6	0.6	308	1	URIC_PNAUV	P34798 canavalia l
290	6	0.6	257	1	LIGF_PSEPA	P30347 pseudomonas	363	6	0.6	308	1	Y325_METJA	O57771 methanococc
291	6	0.6	257	1	Y453_AOUAE	O66740 aquifex aeo	364	6	0.6	309	1	URIC_SOYBN	P04670 glycine max
292	6	0.6	257	1	YHIO_HAEIN	P44901 haemophilus	365	6	0.6	309	1	URID_SOYBN	O54104 glycine max
293	6	0.6	258	1	NCAP_TSWVI	P25999 tomato spot	366	6	0.6	310	1	IFRH_ARATH	P52577 arabidopsis
294	6	0.6	258	1	NCAP_TSWVI	P26000 tomato spot	367	6	0.6	311	1	APE_MOUSE	P08326 mus musculu
295	6	0.6	258	1	NCAP_TSWVI	P26001 tomato spot	368	6	0.6	311	1	SRG6_CAEEL	P54128 caenorhabdl
296	6	0.6	258	1	NCAP_TSWVI	P36294 tomato spot	369	6	0.6	311	1	Y651_METJA	O58067 methanococc
297	6	0.6	258	1	RSFA_BACSG	P39650 bacillus su	370	6	0.6	312	1	APE_RAT	P02650 rattus norv
298	6	0.6	259	1	TRC2_SIRCO	O9z4x0 streptomyc	371	6	0.6	312	1	LIP_VIRCH	P15493 rattus norv
299	6	0.6	259	1	CAH2_MOUSE	P00320 mus musculu	372	6	0.6	313	1	CRIM_MOUSE	O54983 mus musculu
300	6	0.6	259	1	COUL_LEPOS	O91082 i corticotr	373	6	0.6	313	1	NU2M_RHISA	O29817 rhidicephal
301	6	0.6	259	1	CYAA_BACTY	O04470 bacillus th	374	6	0.6	314	1	CRYM_MACFL	O28488 macropus fu
302	6	0.6	261	1	FABI_ECOLI	P29132 escherichia	375	6	0.6	315	1	LPKX_AOUAE	O67572 aquifex aeo
303	6	0.6	261	1	FABI_ECOLI	P44332 haemophilus	376	6	0.6	315	1	RSEB_HAEIN	P44792 haemophilus
304	6	0.6	261	1	FABI_HAEIN	P16657 salmoneilla	377	6	0.6	315	1	SOHC_KLEPN	P37078 klebsiella
305	6	0.6	262	1	PHNC_ECOLI	P16677 escherichia	378	6	0.6	318	1	YORC_ECOLI	O46655 escherichia
306	6	0.6	263	1	MODE_YERPE	P58497 yersinia pe	379	6	0.6	319	1	EXOW_RHIME	P33702 rhizobium m
307	6	0.6	266	1	FLYI_ECOLI	P39174 escherichia	380	6	0.6	321	1	TAZ9_TOBAC	P24804 nicotiana t
308	6	0.6	267	1	BAF_BORPE	O45338 bordetella	381	6	0.6	322	1	ACCI_ORSYA	O40634 oryza sativ
309	6	0.6	267	1	YFNU_ECOLI	P76469 escherichia	382	6	0.6	322	1	ATIZ_VACC	P21114 vaccinia vi
310	6	0.6	268	1	Y489_METJA	O57913 methanococc	383	6	0.6	322	1	YF46_AOUAE	O67500 aquifex aeo
311	6	0.6	269	1	CB2_PHRPA	P20866 physcomitre	384	6	0.6	322	1	YMX7_CAEEL	P34515 caenorhabdl
312	6	0.6	269	1	ESLI_MYCPN	P75333 mycoplasma	385	6	0.6	323	1	ANK5_CYNPY	P70705 cynops pytr
313	6	0.6	272	1	E2F6_MOUSE	O54917 mus musculu	386	6	0.6	323	1	ARGI_RAT	P70724 rattus norv
314	6	0.6	274	1	YE05_MYCTU	P71673 mycobacteri	387	6	0.6	323	1	GM4D_PSEAE	O51366 pseudomonas
315	6	0.6	275	1	PURR_STRPN	O978p1 streptococc	388	6	0.6	323	1	LUKE_STPAU	P31715 staphylococ
316	6	0.6	277	1	PORT_MAYZ	P42057 zea mays (m	389	6	0.6	323	1	YMC2_SCHPO	P22190 schizosacch
317	6	0.6	278	1	CYST_STNP7	P27367 synchococc	390	6	0.6	323	1	YQ34_CAEEL	O09455 caenorhabdl
318	6	0.6	279	1	THET_THEVU	P04072 thermoactin	391	6	0.6	324	1	LEP_SALTY	P23697 salmoneilla
319	6	0.6	281	1	E2F6_HUMAN	O75461 homo sapien	392	6	0.6	325	1	YBEO_STPAU	P07726 staphylococ
320	6	0.6	281	1	RM30_YEAST	P36528 saccharomyc	393	6	0.6	325	1	HLGB_ECOLI	P77324 escherichia
321	6	0.6	281	1	YCXB_ARUGR	P31919 auglena gra	394	6	0.6	328	1	GLK_NEIMA	O93633 nisseria m
322	6	0.6	282	1	KC2C_ARUGR	P40329 arabidopsis	395	6	0.6	329	1	ERG3_SCHPO	O13666 schizosacch
323	6	0.6	282	1	YC80_GUTTH	O78449 guillardia	396	6	0.6	329	1	YIID_ECOLI	P32148 escherichia
324	6	0.6	284	1	CX32_MICUN	P51915 micropogoni	397	6	0.6	332	1	GBLP_DICDI	P46800 dictyostell
325	6	0.6	284	1	PPNK_HELPY	O9zj81 helicobacte	398	6	0.6	333	1	CATH_MOUSE	P49935 mus musculu

399	6	0.6	333	1	CATH_RAT	P00786	rattus norv	472	6	0.6	366	1	YHK0_YEAST	P38772	saccharomyc
400	6	0.6	335	1	CBFC_MOUSE	P70353	mus musculu	473	6	0.6	367	1	DUS1_MOUSE	P28563	mus musculu
401	6	0.6	335	1	CBFC_RAT	P62725	rattus norv	474	6	0.6	367	1	GOX2_ARATH	O91tr9	arabidopsis
402	6	0.6	336	1	G3P2_ANAVA	P34917	anabaena va	475	6	0.6	367	1	YRAM_BACSU	O07931	bacillus su
403	6	0.6	336	1	VG27_BPMDA	P64221	mycobacteri	476	6	0.6	368	1	UCR2_YEAST	P07257	saccharomyc
404	6	0.6	337	1	CLT1_HUMAN	O92771	homo sapien	477	6	0.6	369	1	IDH2_YEAST	P28241	saccharomyc
405	6	0.6	337	1	G3P2_ANASP	P58554	anabaena sp	478	6	0.6	370	1	POPF_ECOLI	P31133	escherichia
406	6	0.6	337	1	G3P_SCHCO	P32638	schizophyll	479	6	0.6	370	1	V2R_BOVIN	P36945	bos taurus
407	6	0.6	337	1	PLSX_HELPR	O24693	helicobacte	480	6	0.6	371	1	4HDB_CLOK1	P38945	clostridium
408	6	0.6	337	1	VGH_BPQ4	P03647	bacteriopho	481	6	0.6	371	1	ISDF_CAMJE	O9pm68	c ispd/ispf
409	6	0.6	337	1	YEB7_YEAST	P39996	saccharomyc	482	6	0.6	372	1	ENO_CHLRE	P31683	chlamydomon
410	6	0.6	338	1	CELF_HSV11	P10237	herpes simp	483	6	0.6	373	1	DDI_MYCSM	O9sgn0	mycobacteri
411	6	0.6	338	1	CELF_HSV19	O07005	herpes simp	484	6	0.6	373	1	MDHM_CHLRE	O42686	chlamydomon
412	6	0.6	339	1	CELF_HSV1K	P03178	herpes simp	485	6	0.6	373	1	TYRA_ECOLI	P07023	escherichia
413	6	0.6	339	1	ADH2_BACST	P42327	bacillus st	486	6	0.6	373	1	TYRA_ERWHE	O02287	erythrina her
414	6	0.6	339	1	CLT1_RAT	O924t8	rattus norv	487	6	0.6	375	1	V375_ASFL5	P26111	african swi
415	6	0.6	339	1	MDH_METFE	P16142	methanother	488	6	0.6	377	1	PGT1_HUMAN	P52610	homo sapien
416	6	0.6	339	1	PLSX_HELPR	O9zmm1	helicobacte	489	6	0.6	377	1	PGT1_RAT	P33610	rattus norv
417	6	0.6	341	1	GPD2_MYCTU	O53761	mycobacteri	490	6	0.6	377	1	TRAF_BACST	P33610	rattus norv
418	6	0.6	341	1	Y011_RICPR	O9zed2	ticketsia	491	6	0.6	379	1	DHPH_SPOUR	O45620	bacillus st
419	6	0.6	342	1	RA51_LYCES	O40134	lycopersico	492	6	0.6	380	1	DP3B_BACHD	P97014	sporosarcin
420	6	0.6	343	1	HRCB_BACHD	O9Kd74	bacillus ha	493	6	0.6	380	1	YVYH_BACSU	O9rcal	bacillus su
421	6	0.6	344	1	HRCB_BACST	O45550	bacillus st	494	6	0.6	382	1	GP44_MOUSE	P39131	bacillus su
422	6	0.6	344	1	YA22_AQUAE	O67134	aquifex aeo	495	6	0.6	383	1	PTC_HUMAN	O922j6	mus musculu
423	6	0.6	345	1	VCAA_BPT7	P19726	bacteriopho	496	6	0.6	383	1	NAPA_ENTHR	P01880	homo sapien
424	6	0.6	346	1	AMBP_MERUN	O62377	metionens un	497	6	0.6	384	1	MTLD_CLOAB	P26235	enterococcu
425	6	0.6	346	1	ST50_YEAST	P25344	saccharomyc	498	6	0.6	385	1	COAT_BPHK7	O65992	clostridium
426	6	0.6	346	1	TRA3_LEPBO	O48514	leptospira	499	6	0.6	385	1	SWY_PYPAB	P49861	bacteriopho
427	6	0.6	347	1	HEM2_CAUCR	P57777	caulobacter	500	6	0.6	386	1	GATM_PIG	O9uy11	pyrococcus
428	6	0.6	347	1	IDH_MEMJA	O58991	methanococc	501	6	0.6	389	1	CBX8_HUMAN	P50441	mus scrofa
429	6	0.6	347	1	UI33_HSV6U	P52380	human herpe	502	6	0.6	392	1	TRB1_METHR	O9nc52	homo sapien
430	6	0.6	347	1	VCAA_BPT3	P19693	bacteriopho	503	6	0.6	393	1	CSP_PLABR	P27696	methanobact
431	6	0.6	348	1	PLSX_SYNY3	P73693	synchocyst	504	6	0.6	393	1	YGIC_HAEIN	P14593	plasmodium
432	6	0.6	349	1	ADH1_ASFPL	P41747	aspergillus	505	6	0.6	394	1	GATR_MOUSE	P44940	haemophilus
433	6	0.6	349	1	DKK3_MOUSE	O9qun9	mus musculu	506	6	0.6	394	1	MEMB_METTR	P27354	methylosinu
434	6	0.6	349	1	OAS1_PIG	O29599	sus scrofa	507	6	0.6	394	1	PGK_BACHD	O9K714	bacillus ha
435	6	0.6	349	1	OPSB_SAIIB	O13092	salimiri bol	508	6	0.6	394	1	PGK_BACSU	P49924	bacillus su
436	6	0.6	349	1	RADA_PYRRU	O74036	pyrococcus	509	6	0.6	394	1	VATC_SCHPO	O9mdw6	schizosacch
437	6	0.6	349	1	VC22_VAVR	P34015	variola vir	510	6	0.6	396	1	ARRS_XENLA	P51477	xenopus lae
438	6	0.6	350	1	DKK3_HUMAN	O9ubp4	homo sapien	511	6	0.6	397	1	CD4_CERTO	O08336	cercarobes
439	6	0.6	350	1	REDD_STRCO	P16922	streptomyce	512	6	0.6	397	1	CD4_ERYPA	O08339	erythrocebu
440	6	0.6	351	1	MODC_HAEIN	P45321	haemophilus	513	6	0.6	397	1	LICH_MOUSE	O920m5	mus musculu
441	6	0.6	351	1	MRAY_BORBU	O44776	borrelia bu	514	6	0.6	398	1	VCAB_BPT7	P19727	bacteriopho
442	6	0.6	352	1	DMPF_PSEBP	P19734	pseudomonas	515	6	0.6	398	1	VE2_HPV42	P27223	human papil
443	6	0.6	352	1	FAS_ANSAN	P36189	anser anser	516	6	0.6	399	1	FTS2_BORBU	P45483	borrelia bu
444	6	0.6	352	1	XVNA_DICTH	O12603	dictyoglonu	517	6	0.6	399	1	METK_BACHD	O9K779	bacillus ha
445	6	0.6	353	1	CCR4_FELCA	P56498	fella silve	518	6	0.6	399	1	RARP_P14HA	P22044	human parai
446	6	0.6	353	1	GB1L_ORYLA	P87383	oryzias lat	519	6	0.6	400	1	DACC_ECOLI	P08506	escherichia
447	6	0.6	353	1	VM17_BORHE	P32777	borrelia he	520	6	0.6	401	1	NOE4_RHIME	P06231	rhizobium m
448	6	0.6	355	1	REF_HAEIN	P45441	haemophilus	521	6	0.6	401	1	PGK_SYNY3	P74421	rhizobium m
449	6	0.6	356	1	PCB1_HUMAN	O15365	homo sapien	522	6	0.6	401	1	PLTC_PSEPU	P36641	pseudomonas
450	6	0.6	356	1	PCB1_RABIT	O19948	oryctolagus	523	6	0.6	402	1	NODE_RHIME	P06230	rhizobium m
451	6	0.6	356	1	RADA_PYPAB	O9V733	pyrococcus	524	6	0.6	402	1	RNDP_PIG	P17560	sus scrofa
452	6	0.6	357	1	MCAL_RHOOP	O84992	rhodococcus	525	6	0.6	403	1	SHBG_MOUSE	P97497	mus musculu
453	6	0.6	357	1	Y180_ARHOC	O80543	arabidopsis	526	6	0.6	404	1	AK1_BACSU	O04795	bacillus su
454	6	0.6	357	1	Y234_LISMO	O48762	listeria mo	527	6	0.6	404	1	TI1BC_HUMAN	P29466	homo sapien
455	6	0.6	357	1	Y266_LISIN	O92741	listeria in	528	6	0.6	405	1	DHMH_METME	O35542	methylophil
456	6	0.6	358	1	DYP3_RAT	O62921	rattus norv	529	6	0.6	405	1	LEU3_YARLI	P18120	yarowia li
457	6	0.6	358	1	PLCA_BRWCH	O47499	erwinia chr	530	6	0.6	406	1	Y4RN_RHISN	P55647	rhizobium s
458	6	0.6	359	1	A2HS_BOVIN	P12763	bos taurus	531	6	0.6	406	1	YDLB_SCHPO	P87126	schizosacch
459	6	0.6	359	1	NHR2_CAEEL	O10902	caenorhabdi	532	6	0.6	407	1	COAE_MYCTU	O06148	mycobacteri
460	6	0.6	359	1	OCD_BRUBA	O59175	bruceella ab	533	6	0.6	409	1	GUB_BACCI	P19254	mycobacteri
461	6	0.6	359	1	PEXC_CAEEL	O19189	caenorhabdi	534	6	0.6	409	1	GONB_RUMAL	P23661	ruminococcu
462	6	0.6	359	1	RFL_CHILU	O9p116	chlamydia t	535	6	0.6	409	1	NER1_MOUSE	O36567	mus musculu
463	6	0.6	359	1	RFL_CHLTR	O84026	chlamydia m	536	6	0.6	409	1	NER1_RAT	O99pw3	rattus norv
464	6	0.6	359	1	YJ72_SYNY3	P74171	synchocyst	537	6	0.6	409	1	SYH12_SCHPO	O43092	schizosacch
465	6	0.6	361	1	GMD1_ARATH	O9enry3	arabidopsis	538	6	0.6	409	1	COAE_MYCTU	O28631	archaeoglob
466	6	0.6	361	1	ODPB_ASCSU	P26269	ascaris suu	539	6	0.6	410	1	COAE_MYCTU	O50178	mycobacteri
467	6	0.6	361	1	V227_FOWPY	O9J508	fowlpox vir	540	6	0.6	410	1	PRSI_METHR	O26824	methanobact
468	6	0.6	362	1	PCB2_MOUSE	O61990	mus musculu	541	6	0.6	410	1	SEX3_MYCTU	O01129	mycobacteri
469	6	0.6	364	1	A2HS_SHEEP	P29701	ovis aries	542	6	0.6	410	1	CLPX_PASMU	O11155	mycobacteri
470	6	0.6	365	1	PCB2_HUMAN	O15366	homo sapien	543	6	0.6	411	1	VGDM_HSVBC	P57981	pasteurella
471	6	0.6	366	1	FLHF_BACSU	O01960	bacillus su	544	6	0.6	411	1		P52370	bovine herp

545	6	0.6	413	1	B2AR_HUMAN	P07550	homo sapien	618	6	0.6	453	1	SYS_ARCFU	O28244	archaeoglob
546	6	0.6	413	1	FTSA_BORBU	O44774	borrelia bu	619	6	0.6	455	1	MENE_SALTY	P37418	salmonella
547	6	0.6	413	1	HEM1_CORBU	P47846	coxiella hu	620	6	0.6	456	1	PROA_YEAST	P54885	saccharomyc
548	6	0.6	413	1	TAPC_AERHY	P45733	aeromonas h	621	6	0.6	457	1	CD4_SAISC	O29037	salimiri scl
549	6	0.6	414	1	GAG2_DROME	P20828	drosophila	622	6	0.6	457	1	FTS2_PORGI	O08466	porphyromon
550	6	0.6	414	1	YGJ0_HAEIN	P45264	haemophilus	623	6	0.6	457	1	RADA_BACDH	O9k991	baacillus ha
551	6	0.6	415	1	B2AR_CANFA	P54833	canis fam1l	624	6	0.6	457	1	RADA_LISIN	O92f42	listeria in
552	6	0.6	415	1	B2AR_MACMU	O28509	macaca mula	625	6	0.6	457	1	RADA_LISMO	O48791	listeria mo
553	6	0.6	415	1	YAE2_YEAST	P39727	saccharomyc	626	6	0.6	458	1	CBFC_HUMAN	O13952	homo sapien
554	6	0.6	416	1	CCA_HAEIN	P45269	haemophilus	627	6	0.6	458	1	CD4_CERAE	O08338	cercopithec
555	6	0.6	417	1	HS47_HUMAN	P29043	homo sapien	628	6	0.6	458	1	CD4_HUMAN	P01730	homo sapien
556	6	0.6	417	1	HS47_DROME	P19334	mus musculu	629	6	0.6	458	1	CD4_MACPA	P79185	macaca fasc
557	6	0.6	417	1	HS47_MOUSE	P29457	rattus norv	630	6	0.6	458	1	CD4_MACPU	P79185	macaca fusc
558	6	0.6	418	1	B2AR_FELCA	O9ts55	felis silve	631	6	0.6	458	1	CD4_MACMU	P16003	macaca mula
559	6	0.6	418	1	CBP2_HUMAN	P50454	homo sapien	632	6	0.6	458	1	CD4_MACNE	O08340	macaca neme
560	6	0.6	419	1	PGK_TREPA	O83549	treponema p	633	6	0.6	458	1	CD4_PANTR	P16004	pan troglod
561	6	0.6	420	1	KC2A_THEPA	P28547	thellieria p	634	6	0.6	458	1	ODP2_NECCR	P20285	neutrospora
562	6	0.6	420	1	VIT3_DROME	O66607	drosophila	635	6	0.6	458	1	RADA_BACSU	P37572	baacillus su
563	6	0.6	420	1	VIT3_SYNY3	O5874	synecocyst	636	6	0.6	458	1	RBL2_RHOCA	P50922	rhodobacter
564	6	0.6	421	1	HEMA_CVBF	P33468	bovine coro	637	6	0.6	459	1	ND4M_CHICK	P18939	gallus gall
565	6	0.6	421	1	RL4_RAT	P50878	rattus norv	638	6	0.6	460	1	SELA_PASMU	O9kx66	pasteurella
566	6	0.6	423	1	SHP1_YEAST	P34223	saccharomyc	639	6	0.6	460	1	VS1K_ACLSP	P27739	apple chlor
567	6	0.6	424	1	HEMA_CVBLV	P31613	bovine coro	640	6	0.6	461	1	MTS1_SALIN	P09795	salmonella
568	6	0.6	424	1	HEMA_CVBM	P15776	bovine coro	641	6	0.6	461	1	SELA_HAEIN	P43910	haemophilus
569	6	0.6	424	1	HEMA_CVBO	P24351	bovine coro	642	6	0.6	461	1	TBG1_EUPOC	P34786	euplotes oc
570	6	0.6	424	1	HEMA_CVHOC	P30215	human coron	643	6	0.6	461	1	TBG2_EUPOC	P90548	euplotes ac
571	6	0.6	424	1	VG46_BPPEI	P25131	bacterioph	644	6	0.6	461	1	TBG_EUPAE	P54402	euplotes ae
572	6	0.6	425	1	CABL_CAREL	O93249	caenorhabdi	645	6	0.6	462	1	DPD2_SCHPO	P87334	schizosach
573	6	0.6	426	1	TOLB_CHLMU	O9p361	chlamydia m	646	6	0.6	462	1	TBG1_EUPCR	P54403	euplotes cr
574	6	0.6	427	1	ACEA_BACHD	O9k910	baacillus ha	647	6	0.6	463	1	SAHH_RHOCA	P28183	rhodobacter
575	6	0.6	427	1	RL4_HUMAN	P36578	homo sapien	648	6	0.6	464	1	DNAA_TREPA	O83047	treponema p
576	6	0.6	428	1	AROA_EDMIC	O94K12	edwardstiell	649	6	0.6	465	1	CHLN_MARPO	P26156	marichantia
577	6	0.6	428	1	GFAP_BOVIN	O28115	bos taurus	650	6	0.6	465	1	COBO_ACIICA	O43989	acinetobact
578	6	0.6	428	1	CSP_PLAMA	P13815	plasmodium	651	6	0.6	465	1	HYTN_BRAJA	P19922	bradyrhizob
579	6	0.6	429	1	SUN_ECOLI	P36929	escherichia	652	6	0.6	466	1	CYP8_CAREL	P32016	caenorhabdi
580	6	0.6	430	1	YHT2_YEAST	P38836	saccharomyc	653	6	0.6	466	1	IMP2_DROME	P29661	drosophila
581	6	0.6	431	1	ENO_ECOLI	P08344	escherichia	654	6	0.6	466	1	SERA_SCHPO	P87228	schizosach
582	6	0.6	431	1	TOLB_CHLTR	O84604	chlamydia t	655	6	0.6	467	1	EUTE_ECOLI	P77445	escherichia
583	6	0.6	432	1	AROA_HARSO	P52310	haemophilus	656	6	0.6	467	1	EVTE_SALTY	P41793	salmonella
584	6	0.6	432	1	AROA_PASHA	P54220	pasteurella	657	6	0.6	468	1	ATPB_MYCPU	O98466	mycoplasma
585	6	0.6	432	1	GFAP_HUMAN	P14136	homo sapien	658	6	0.6	469	1	PPAF_ARATH	O83924	arabidopsis
586	6	0.6	432	1	KGTP_ECOLI	P17448	escherichia	659	6	0.6	470	1	LEU2_AZOVI	P61675	azotobacter
587	6	0.6	432	1	PROA_DEIRA	O9t439	deinococcus	660	6	0.6	470	1	SELA_MOOTH	O33277	moreella th
588	6	0.6	432	1	TIG_HAEIN	P44837	haemophilus	661	6	0.6	470	1	UDPG_ARATH	P57751	arabidopsis
589	6	0.6	432	1	YGB1_YEAST	P48326	saccharomyc	662	6	0.6	471	1	GATH_SCHPO	O94509	schizosach
590	6	0.6	433	1	COBB_PSEDE	P21632	pseudomonas	663	6	0.6	473	1	MAIR_YEAST	P53338	saccharomyc
591	6	0.6	433	1	PGKH_SPIOI	P29409	spinaclia ol	664	6	0.6	474	1	GATB_HELPU	O92116	helicobacte
592	6	0.6	433	1	SAH1_XENLA	P11893	xenopus lae	665	6	0.6	474	1	GATB_HELPU	O25372	helicobacte
593	6	0.6	433	1	VCAB_BPT3	P19728	bacterioph	666	6	0.6	474	1	PRTP_MOUSE	P16675	mus musculu
594	6	0.6	435	1	ASPG_MERTH	O26802	methanobact	667	6	0.6	475	1	PURL_MERTH	P26742	methanobact
595	6	0.6	435	1	GLSF_HORVU	O08258	hordewm vul	668	6	0.6	475	1	GUNA_CIOCE	P17901	clostridium
596	6	0.6	435	1	YOD7_CAREL	O09267	caenorhabdi	669	6	0.6	475	1	MMR_STRO	P11545	streptomyce
597	6	0.6	438	1	Y281_HAEIN	P44610	haemophilus	670	6	0.6	476	1	PCCB_RHOER	O06101	rhodococcus
598	6	0.6	439	1	G6PD_CHLTR	O84188	chlamydia t	671	6	0.6	476	1	UDPG_SOLTU	P19559	solanum tub
599	6	0.6	439	1	VIT1_DROME	P02843	drosophila	672	6	0.6	477	1	DLDH_HAEIN	P33784	haemophilus
600	6	0.6	441	1	DNB2_ADEG1	O64759	avian adeno	673	6	0.6	477	1	GIGA_ECOLI	P08333	escherichia
601	6	0.6	442	1	TIG_CHLTR	O84713	chlamydia t	674	6	0.6	477	1	HLVD_ACTAC	P18790	actinobacil
602	6	0.6	443	1	FLIT_AQUAE	O67531	aquifex aeo	675	6	0.6	477	1	VNUC_DHYI	P11468	dhori virus
603	6	0.6	443	1	LCPB_LEIME	P36400	leishmania	676	6	0.6	478	1	HLV4_ECOLI	P09966	escherichia
604	6	0.6	444	1	CYS2_LEIPI	O05094	leishmania	677	6	0.6	478	1	HLVD_ECOLI	P67329	escherichia
605	6	0.6	445	1	Y148_MYCPN	P75584	mycoplasma	678	6	0.6	478	1	NRFA_ECOLI	P32050	escherichia
606	6	0.6	445	1	YR82_MYCLE	O32965	mycobacteri	679	6	0.6	478	1	PGKH_ARATH	P50318	arabidopsis
607	6	0.6	446	1	TBB5_CHICK	P09653	gallus gall	680	6	0.6	478	1	YARR_HRTSN	P55641	rhizobium s
608	6	0.6	447	1	PROA_METAC	O9hba1	methanosarc	681	6	0.6	480	1	KLFA_HUMAN	O13118	homo sapien
609	6	0.6	447	1	SR34_MYCY	O01442	mycoplasma	682	6	0.6	480	1	KLFA_RAT	O08786	rattus norv
610	6	0.6	449	1	VNS5_INSVN	O01811	impatiens n	683	6	0.6	480	1	PGKH_MHEAT	P12782	triticum ae
611	6	0.6	450	1	GSHR_STRTR	O60151	streptococc	684	6	0.6	480	1	PTSB_STAXY	P51180	staphylococ
612	6	0.6	451	1	MENE_ECOLI	P37333	escherichia	685	6	0.6	480	1	UCRI_BOVIN	P42469	bos taurus
613	6	0.6	452	1	TRME_PASMU	O9C161	pasteurella	686	6	0.6	481	1	ATPB_STIAY	P42469	stigmatis b
614	6	0.6	453	1	ALN_ECOLI	P77671	escherichia	687	6	0.6	481	1	DPOL_HPBAY	P03318	hepatitis b
615	6	0.6	453	1	PH4H_MOUSE	P16331	mus musculu	688	6	0.6	481	1	PGKH_TBOAC	O42961	nicotiana t
616	6	0.6	453	1	RADA_PSEAE	P69563	pseudomonas	689	6	0.6	483	1	ENP1_YEAST	P38333	saccharomyc
617	6	0.6	453	1	RADA_STRPY	O9a1k1	streptococc	690	6	0.6	483	1	ODP2_SCHPO	O59816	schizosach

691	6	0.6	483	1	PREG_NEUCR	006112	neurospora	764	6	0.6	523	1	YHUS_ECOLI	P37657	eschlerichia
692	6	0.6	483	1	PVAB_ECOLI	P32154	eschlerichia	765	6	0.6	524	1	SAP_HUMAN	P07602	h proactiva
693	6	0.6	483	1	Y045_MYCCE	P47591	mycoplasma	766	6	0.6	525	1	GUAA_PSEAE	O9hm6	pseudomonas
694	6	0.6	484	1	DNB2_ADEI2	P04498	human adeno	767	6	0.6	525	1	SAP_BOVIN	P05779	p proactiva
695	6	0.6	484	1	GLGA_VIRB2	O9krb6	vibrio chol	768	6	0.6	525	1	SP1_PARFA	O05508	ratrobacter
696	6	0.6	484	1	HEMZ_HORVU	P42045	hordeum vul	769	6	0.6	527	1	KCCG_RAT	P11730	rattus norv
697	6	0.6	484	1	SYT1_AERPE	O9ywo0	aeropyrum p	770	6	0.6	527	1	TCPB_SCHPO	Q10147	schlzosach
698	6	0.6	484	1	XYLB_ECOLI	P09099	eschlerichia	771	6	0.6	527	1	TCPD_SCHPO	P05999	schlzosach
699	6	0.6	484	1	Y130_MYCCE	P47376	mycoplasma	772	6	0.6	529	1	RADA_PYRHO	O58081	pyrococcus
700	6	0.6	485	1	DH4A_HUMAN	P51468	homo sapien	773	6	0.6	529	1	YEJF_ECOLI	P33916	eschlerichia
701	6	0.6	485	1	HXKA_YEAST	P04806	saccharomyc	774	6	0.6	530	1	FM03_RABIT	P32417	oryctolagus
702	6	0.6	485	1	Z165_HUMAN	P49910	homo sapien	775	6	0.6	530	1	PPCK_STAMM	P32417	oryctolagus
703	6	0.6	486	1	CISX_YEAST	P43635	saccharomyc	776	6	0.6	530	1	VC04_SPVKA	P51065	staphylococ
704	6	0.6	486	1	YAMB_SCHPO	Q10063	schlzosach	777	6	0.6	531	1	FM03_HUMAN	P32228	swinepox v1
705	6	0.6	486	1	YDBH_SCHPO	Q10367	schlzosach	778	6	0.6	531	1	PGL1_RALSO	P31513	homo sapien
706	6	0.6	488	1	NU2M_OENBE	P93401	oenothera b	779	6	0.6	531	1	RO60_DEIRA	P58598	ralstonia s
707	6	0.6	488	1	SECY_HALVO	O977v3	halobacteri	780	6	0.6	532	1	TYRO_RANNT	O9rw88	deinococcus
708	6	0.6	489	1	NU2M_MARPO	P26846	marchantia	781	6	0.6	533	1	FCY2_YEAST	O04604	rana nigrom
709	6	0.6	490	1	DNEP_YEAST	P38821	saccharomyc	782	6	0.6	534	1	FM03_YEAST	P17064	saccharomyc
710	6	0.6	490	1	V094_HSV6U	Q00683	human herpe	783	6	0.6	535	1	XYNB_BACPU	P97501	mus musculu
711	6	0.6	491	1	CD5_RAT	P51882	rattus norv	784	6	0.6	536	1	ARP_ARATH	P07129	bacillus pu
712	6	0.6	491	1	HSF1_CHICK	P38529	gallus gall	785	6	0.6	536	1	VNH1_CAEEL	P45951	arabidopsis
713	6	0.6	491	1	VIEL_HCMVA	P13302	human cytom	786	6	0.6	540	1	CH61_STROO	P32739	caenorhabdi
714	6	0.6	491	1	VIEL_HCMVA	P03169	human cytom	787	6	0.6	540	1	CH61_STRLI	P40171	streptomyce
715	6	0.6	492	1	YNM4_YEAST	P53919	human cytom	788	6	0.6	541	1	TCPE_HUMAN	O33659	streptomyce
716	6	0.6	493	1	VE2_HPV19	P36786	human papil	789	6	0.6	541	1	TCPE_MOUSE	P48643	homo sapien
717	6	0.6	494	1	CD5_MOUSE	P133799	mus musculu	790	6	0.6	542	1	Y106_YEAST	P80316	mus musculu
718	6	0.6	494	1	COBQ_MYCTU	O53677	mycobacteri	791	6	0.6	543	1	CP1B_MOUSE	P40445	saccharomyc
719	6	0.6	494	1	SECY_AERPE	O9yvd0	aeropyrum p	792	6	0.6	544	1	TCPB_MOUSE	O64429	mus musculu
720	6	0.6	494	1	SPR4_HUMAN	O08170	homo sapien	793	6	0.6	546	1	CH61_RHOSH	P80313	mus musculu
721	6	0.6	494	1	VE2_HPV12	P36782	human papil	794	6	0.6	546	1	Y341_HUMAN	P20110	rhodopacter
722	6	0.6	495	1	AMPA_PSEAE	O68622	pseudomonas	795	6	0.6	548	1	ACEA_NEUCR	O15049	homo sapien
723	6	0.6	495	1	CD5_HUMAN	P06127	homo sapien	796	6	0.6	548	1	ERF_HUMAN	P28299	neurospora
724	6	0.6	496	1	ADA_MYCTU	Q10630	mycobacteri	797	6	0.6	550	1	YERC_CAEEL	P50548	homo sapien
725	6	0.6	496	1	GLPK_BACSU	P18157	thermus aqu	798	6	0.6	551	1	FERF_MOUSE	Q21802	caenorhabdi
726	6	0.6	496	1	GLPK_THEAO	O9xx53	thermus aqu	799	6	0.6	552	1	FERF_YEAST	P70459	mus musculu
727	6	0.6	496	1	GLPK_THEPL	O66131	thermus aqu	800	6	0.6	553	1	NADB_SYNY3	P40988	saccharomyc
728	6	0.6	497	1	C71C_ARATH	O49340	arabidopsis	801	6	0.6	553	1	PPCK_CANAL	P74562	synecocyst
729	6	0.6	497	1	C71D_ARATH	O49342	arabidopsis	802	6	0.6	555	1	PDRA_ECOLI	Q13434	candida alb
730	6	0.6	497	1	C71G_ARATH	O9f1m6	arabidopsis	803	6	0.6	556	1	TCP1_MOUSE	Q47208	eschlerichia
731	6	0.6	499	1	C72L_ARATH	O9f1m2	arabidopsis	804	6	0.6	556	1	TCP2_MOUSE	P11984	mus musculu
732	6	0.6	500	1	C72M_ARATH	O9f1m1	arabidopsis	805	6	0.6	556	1	TCPA_CRIGR	P11983	mus musculu
733	6	0.6	500	1	YR31_YEAST	P43601	saccharomyc	806	6	0.6	556	1	TCPA_HUMAN	P18279	cricetus
734	6	0.6	501	1	ACHB_HUMAN	P11230	homo sapien	807	6	0.6	556	1	TCPA_RAT	P17987	homo sapien
735	6	0.6	501	1	ACHB_MOUSE	P09690	mus musculu	808	6	0.6	557	1	TCPA_MONDO	P28480	rattus norv
736	6	0.6	501	1	ACHB_RAT	P25109	rattus norv	809	6	0.6	557	1	KSD1_ECOLI	O09x06	monodelphis
737	6	0.6	502	1	C718_MENPI	Q42716	mentha pipe	810	6	0.6	558	1	KSD5_ECOLI	Q03961	eschlerichia
738	6	0.6	502	1	VE2_HPV25	P36787	human papil	811	6	0.6	558	1	PILE_NEIGO	P42213	eschlerichia
739	6	0.6	503	1	REL_AVIRE	P01126	avian retic	812	6	0.6	558	1	TCPB_TETPY	P37094	neisseria g
740	6	0.6	505	1	ACHB_BOVIN	P04758	bos taurus	813	6	0.6	561	1	HLTB_PROMI	P54409	tetratymena
741	6	0.6	506	1	CP44_RABIT	P10611	oryctolagus	814	6	0.6	562	1	4CL2_SOYBN	P16465	proteus mir
742	6	0.6	507	1	G6PD_CHLMO	O9pkk8	chlamydia m	815	6	0.6	563	1	HNM1_YEAST	P31687	glycine max
743	6	0.6	507	1	PSBB_SYNY3	P05429	synecocyst	816	6	0.6	563	1	SRC8_CHICK	P19807	saccharomyc
744	6	0.6	509	1	ATPO_NICPL	P05495	nicotiana p	817	6	0.6	564	1	DPY4_RAT	Q01406	gallus gall
745	6	0.6	509	1	CPPT7_PIG	P19100	sus scrofa	818	6	0.6	564	1	YC62_AQUAE	O62951	rattus norv
746	6	0.6	509	1	VE2_HPV36	P50809	human papil	819	6	0.6	566	1	CRT1_ARATH	O67301	agulfex aeo
747	6	0.6	509	1	YF08_MYCPN	P75278	mycoplasma	820	6	0.6	568	1	TAPB_AERHY	Q07356	arabidopsis
748	6	0.6	510	1	CP46_RABIT	P14808	oryctolagus	821	6	0.6	569	1	MAIT_CANAL	P45792	aeromonas h
749	6	0.6	510	1	ENP1_MOUSE	P55772	mus musculu	822	6	0.6	569	1	TACT_HUMAN	Q02512	candida alb
750	6	0.6	511	1	CP45_RABIT	P14579	oryctolagus	823	6	0.6	570	1	DPY3_HUMAN	P40200	homo sapien
751	6	0.6	511	1	CP47_RABIT	P14581	oryctolagus	824	6	0.6	570	1	DPY3_MOUSE	O14195	homo sapien
752	6	0.6	511	1	ENP1_RAT	P97687	rattus norv	825	6	0.6	570	1	SM44_CAEEL	O62188	mus musculu
753	6	0.6	512	1	ALDB_ECOLI	P37685	eschlerichia	826	6	0.6	571	1	DPY3_XENLA	P45897	caenorhabdi
754	6	0.6	512	1	C7C1_ARATH	O64366	arabidopsis	827	6	0.6	571	1	PAL4_POPOKI	O13022	xenopus lae
755	6	0.6	513	1	CADA_DROMT	O18596	drosophila	828	6	0.6	572	1	DPY1_HUMAN	Q04910	populus kit
756	6	0.6	514	1	G6PD_ARATH	Q43727	arabidopsis	829	6	0.6	572	1	DPY1_MOUSE	O14194	homo sapien
757	6	0.6	517	1	YXWL_CAEEL	O20806	caenorhabdi	830	6	0.6	572	1	DPY1_RAT	P97427	mus musculu
758	6	0.6	518	1	GLG2_ARATH	P55230	caenorhabdi	831	6	0.6	572	1	DPY2_BOVIN	O62950	rattus norv
759	6	0.6	519	1	ATPB_NEUCR	P23704	neurospora	832	6	0.6	572	1	DPY2_CHICK	O02675	bos taurus
760	6	0.6	519	1	RHO_TREPA	O83881	treponema p	833	6	0.6	572	1	DPY2_HUMAN	Q09635	gallus gall
761	6	0.6	520	1	CZCB_ALCEU	P13510	alcaligenes	834	6	0.6	572	1	DPY2_MOUSE	O16555	homo sapien
762	6	0.6	520	1	CZCB_ALGSP	P94176	alcaligenes	835	6	0.6	572	1	DPY2_MOUSE	O08553	mus musculu
763	6	0.6	522	1	KRFL_RAT	O63369	rattus norv	836	6	0.6	572	1	DPY4_HUMAN	P47942	rattus norv
														O14531	homo sapien

837	6	0.6	572	1	DPY4_MOUSE	035098	mus musculus	910	6	0.6	643	1	SYRM_YEAST	P38714	saccharomyc
838	6	0.6	573	1	MDL3_PRUSE	P52707	prunus sero	911	6	0.6	644	1	DXS_CHLCP	O92619	chlamydia p
839	6	0.6	573	1	TLPC_BACSU	P39209	bacillus su	912	6	0.6	644	1	VP74_NPYCP	O10365	oryzaria pseu
840	6	0.6	575	1	IL10R_MOUSE	O61727	mus musculus	913	6	0.6	645	1	VP74_NPYCF	P34053	chortistoneu
841	6	0.6	575	1	IL10R_MOUSE	O02117	lactococcus	914	6	0.6	645	1	VP74_NPYCF	O26542	methanobact
842	6	0.6	575	1	YWA_BACSU	P45861	bacillus su	915	6	0.6	646	1	UVRE_MERTH	O59941	streplococc
843	6	0.6	576	1	CSG_METVO	O50833	methanococc	916	6	0.6	647	1	PARE_STRPN	O16643	homo saplen
844	6	0.6	577	1	CHVG_RHIME	P72292	rhizobium m	917	6	0.6	649	1	DREB_HUMAN	O95848	homo saplen
845	6	0.6	577	1	CST2_HUMAN	P33240	homo saplen	918	6	0.6	652	1	C3B8_BACTU	O06117	bacillus th
846	6	0.6	578	1	YCS2_BACSU	P42967	bacillus su	919	6	0.6	652	1	RPSD_CAUCR	P52324	caulobacter
847	6	0.6	580	1	ATPU_YEAST	P22136	saccharomyc	920	6	0.6	652	1	T3WO_SALTY	P48814	salmonella
848	6	0.6	580	1	GGT_ECOLI	P18956	escherichia	921	6	0.6	653	1	GP11_SCHPO	O14337	schizosacch
849	6	0.6	583	1	ASNS_ARATH	P49078	arabidopsis	922	6	0.6	653	1	MALO_PYRO	O32450	pyrococcus
850	6	0.6	583	1	GTBI_MOUSE	O08582	mus musculus	923	6	0.6	657	1	SGV1_YEAST	P32329	saccharomyc
851	6	0.6	583	1	ML1X_MOUSE	O08445	mus musculus	924	6	0.6	657	1	SIM2_MOUSE	O61079	mus musculus
852	6	0.6	584	1	GTBI_HUMAN	O00178	homo saplen	925	6	0.6	657	1	UVRE_MYCPU	P58174	mycoplasma
853	6	0.6	584	1	LMRA_LACLA	O9GHI8	lactococcus	926	6	0.6	658	1	ABD3_RAT	P16970	rattus norv
854	6	0.6	584	1	LMRA_LACLA	P97046	lactococcus	927	6	0.6	658	1	VG18_BPT4	P13332	bacterioph
855	6	0.6	585	1	ASNS_BRAOL	P49091	brassica ol	928	6	0.6	659	1	ABD3_HUMAN	P28288	homo saplen
856	6	0.6	589	1	ASNS_ASPOF	P31752	asparagus o	929	6	0.6	659	1	ABD3_MOUSE	P55056	mus musculus
857	6	0.6	589	1	SYER_DROME	O9VCS5	drosofila	930	6	0.6	659	1	C3BA_BACTO	P17969	bacillus th
858	6	0.6	590	1	VG29_BPT4	P13337	bacterioph	931	6	0.6	662	1	YME1_SCHMA	P46508	schistosoma
859	6	0.6	591	1	GBP2_HUMAN	P32456	homo saplen	932	6	0.6	665	1	FTH2_STYX3	P73179	synecocyst
860	6	0.6	595	1	TRPE_ARATH	P32068	arabidopsis	933	6	0.6	667	1	SIM2_HUMAN	O14150	homo saplen
861	6	0.6	596	1	FRDA_SHEFR	O02469	shewanella	934	6	0.6	670	1	CACP_YEAST	P32796	saccharomyc
862	6	0.6	598	1	ABBI_RAT	P46933	rattus norv	935	6	0.6	671	1	NOG1_ARATH	O9C618	arabidopsis
863	6	0.6	598	1	HS7L_SHRY	P37092	sugar beet	936	6	0.6	671	1	P5CS_YGCAC	P32296	v delta 1-p
864	6	0.6	600	1	PFRI_TRYBB	P22225	trypanosoma	937	6	0.6	672	1	ABG8_RAT	P58488	rattus norv
865	6	0.6	600	1	Y719_METJA	O38129	methanococ	938	6	0.6	674	1	RPSD_RHOCA	P64600	rhodobacter
866	6	0.6	603	1	FTSH_CYAME	O9C139	cyanidiosch	939	6	0.6	675	1	DNLJ_BUCAI	P57172	buchnera ap
867	6	0.6	603	1	LCB2_SCHPO	O09925	schizosacch	940	6	0.6	676	1	EXL1_HUMAN	O92395	homo saplen
868	6	0.6	603	1	NMBL_MOUSE	O08919	mus musculus	941	6	0.6	676	1	EXL1B_HUMAN	O15296	homo saplen
869	6	0.6	604	1	KRAA_RAT	P14056	rattus norv	942	6	0.6	679	1	MTOL_YEAST	P53070	saccharomyc
870	6	0.6	605	1	RTK2_GEOCY	O9WXX5	t glucosami	943	6	0.6	679	1	SYGB_BACSU	P54381	bacillus su
871	6	0.6	605	1	RTK2_GEOCY	P42159	geodia cydo	944	6	0.6	684	1	RPSD_RHIME	O59753	rhizobium m
872	6	0.6	605	1	SYT_MERTH	O27504	methanobact	945	6	0.6	689	1	SYGB_PASMU	P57905	pasturella
873	6	0.6	606	1	KRAA_HUMAN	P10338	homo saplen	946	6	0.6	690	1	CHVG_AGR75	O07737	agrobacteri
874	6	0.6	606	1	KRAA_PIG	O19004	sus scrofa	947	6	0.6	692	1	AKAB_HUMAN	O43823	homo saplen
875	6	0.6	608	1	AFAM_RAT	P36953	rattus norv	948	6	0.6	692	1	DPB2_YEAST	P24482	saccharomyc
876	6	0.6	608	1	EDD_HELPJ	O9ZKX3	helicobacte	949	6	0.6	692	1	Y650_METJA	O57852	methanococ
877	6	0.6	608	1	EDD_HELPJ	P56111	helicobacte	950	6	0.6	701	1	HS90_PODAN	O43109	podospira a
878	6	0.6	609	1	NMBL_HUMAN	O9Y670	homo saplen	951	6	0.6	702	1	SPE1_ARATH	O9S164	arabidopsis
879	6	0.6	610	1	CDP1_ARATH	O06850	arabidopsis	952	6	0.6	703	1	DD35_HUMAN	O9H515	homo saplen
880	6	0.6	611	1	VIB3_MCMVS	P29883	murine cyto	953	6	0.6	703	1	Y411_RHISN	P55492	rhizobium s
881	6	0.6	615	1	SECD_ECOLI	P19673	escherichia	954	6	0.6	704	1	HPS_MOUSE	O08993	mus musculus
882	6	0.6	615	1	SECD_SALCH	O9ZF88	salmonella	955	6	0.6	707	1	DCOR_LETDO	P27116	leishmania
883	6	0.6	618	1	MBHL_RHOGE	P17632	rhodocyclus	956	6	0.6	707	1	DREB_RAT	O07266	rattus norv
884	6	0.6	619	1	PPCK_HAECO	P29130	haemophilus	957	6	0.6	707	1	KPC2_CAEEL	P34885	caenorhabdi
885	6	0.6	619	1	RECO_HAEIN	P71359	haemophilus	958	6	0.6	708	1	ABBI_MOUSE	O9QX11	mus musculu
886	6	0.6	619	1	SR68_HUMAN	O9UHB9	homo saplen	959	6	0.6	710	1	ABBI_HUMAN	O00213	homo saplen
887	6	0.6	620	1	VLCG_RAT	P97554	rattus norv	960	6	0.6	711	1	SPE2_ARATH	O23141	arabidopsis
888	6	0.6	621	1	MX_ONCMY	O91152	oncorhynch	961	6	0.6	712	1	CDGT_BAC3	P09131	bacillus sp
889	6	0.6	621	1	TRPX_ARATH	P32069	arabidopsis	962	6	0.6	713	1	CDGT_BAC3	P05618	bacillus sp
890	6	0.6	622	1	MAK_MOUSE	O04859	mus musculus	963	6	0.6	714	1	YJL2_YEAST	P47025	saccharomyc
891	6	0.6	622	1	MAK_RAT	P20793	rattus norv	964	6	0.6	716	1	COAT_MOUSE	P03137	murine min
892	6	0.6	622	1	SR68_CANFA	O00004	canis fami	965	6	0.6	716	1	P5CS_ORYSA	O04226	o delta 1-p
893	6	0.6	623	1	CA44_RABIT	P55787	oryctolagus	966	6	0.6	717	1	P5CL_ARATH	P54887	a delta 1-p
894	6	0.6	623	1	MAK_HUMAN	P20794	homo saplen	967	6	0.6	717	1	P5CS_ACRCH	O04015	a delta 1-p
895	6	0.6	624	1	AMTG_ARKAD	P42042	arxula aden	968	6	0.6	717	1	P5CS_LYCES	O96460	l delta 1-p
896	6	0.6	625	1	AMVG_NEUCR	P14804	neutrospora	969	6	0.6	718	1	COAT_MUMIM	P07302	murine min
897	6	0.6	626	1	HTPG_ACTYAC	P54669	actinobacil	970	6	0.6	718	1	KHL4_HUMAN	O9C012	homo saplen
898	6	0.6	626	1	MAG_HUMAN	P20916	homo saplen	971	6	0.6	719	1	GROU_DROME	P16371	drosophila
899	6	0.6	626	1	MAG_MOUSE	P20917	mus musculus	972	6	0.6	722	1	COAT_MEVA	P27437	mink enteri
900	6	0.6	626	1	MAG_RAT	P07722	rattus norv	973	6	0.6	722	1	COAT_PAVCH	P04863	canine parv
901	6	0.6	628	1	SNXI_HUMAN	O961F0	homo saplen	974	6	0.6	722	1	COAT_PAVCH	P03136	hamster par
902	6	0.6	632	1	CSX1_SCHPO	O13759	schizosacch	975	6	0.6	722	1	FLID_TREPA	O83842	treponema p
903	6	0.6	632	1	RECO_PASMU	O9C121	pasteurella	976	6	0.6	725	1	ADDB_MOUSE	O9QY8	mus musculus
904	6	0.6	633	1	ROR_HUMAN	O43330	homo saplen	977	6	0.6	725	1	ADDB_RAT	O05744	rattus norv
905	6	0.6	634	1	YC36_METJA	O58663	methanococ	978	6	0.6	725	1	TAP1_RAT	P63710	rattus norv
906	6	0.6	638	1	KNCO_YEAST	P53974	saccharomyc	979	6	0.6	725	1	YJ76_YEAST	P47144	saccharomyc
907	6	0.6	638	1	NOS2_PSEST	P19573	pseudomonas	980	6	0.6	726	1	ADDB_HUMAN	P35612	homo saplen
908	6	0.6	638	1	YOGS_BACSU	P54456	bacillus su	981	6	0.6	726	1	P5C2_ARATH	P54888	a delta 1-p
909	6	0.6	641	1	MTHS_SCHPO	O74937	schizosacch	982	6	0.6	727	1	COAT_FPV	P04864	feline panl

PKC_BACME	STANDARD	PRT	394 AA
983	0.6	727	1 COAT_FPV19
984	0.6	727	1 COAT_PAVCB
985	0.6	729	1 COAT_PAVPB
986	0.6	729	1 COAT_PAVPK
987	0.6	729	1 COAT_PAVPN
988	0.6	729	1 DHTR_METME
989	0.6	734	1 PSAB_MARPO
990	0.6	736	1 PRXL_HUMAN
991	0.6	737	1 COAT_PAVCD
992	0.6	737	1 PRXL_MOUSE
993	0.6	738	1 PLO3_HUMAN
994	0.6	739	1 DTD_HUMAN
995	0.6	739	1 DTD_MOUSE
996	0.6	739	1 DTD_RAT
997	0.6	740	1 Y050_HUMAN
998	0.6	745	1 PNP_RICPR
999	0.6	746	1 RNC3_SCHPO
1000	0.6	748	1 COAT_PAVCN

ALIGNMENTS

RESULT 1

PKC_BACME STANDARD; PRT: 394 AA.

AC P24269;

DT 01-MAR-1992 (Rel. 21, Created)

DT 01-MAR-1992 (Rel. 21, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Phosphoglycerate kinase (EC 2.7.2.3).

GN PGK.

OS Bacteria; megaterium.

OC Bacteria; Firmicutes; Bacillus/Clostridium group;

OC Bacillus/Staphylococcus group; Bacillus.

OX NCBI_TaxID=1404;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-DSM 319.

RA MEDLINE-91057129; PubMed-2123031;

RA Schlaepfer B.S., Brantant C., Brantant G., Zuber H.;

RT "Nucleotide sequence of the phosphoglycerate kinase gene from

RT Bacillus megaterium";

RL Nucleic Acids Res. 18:6423-6423(1990).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN-DSM 319.

RX MEDLINE-93083995; PubMed-1452037;

RA Schlaepfer B.S., Zuber H.;

RT "Cloning and sequencing of the genes encoding

RT glyceraldhyde-3-phosphate dehydrogenase, phosphoglycerate kinase and

RT triosephosphate isomerase (gap operon) from mesophilic Bacillus

RT megaterium: comparison with corresponding sequences from thermophilic

RT Bacillus stearothermophilus.";

RL Gene 122:53-62(1992)

RT Bacterial Activity: ATP + 3-phospho-D-glycerate = ADP + 3-

RT CATALYTIC ACTIVITY: ATP + 3-phospho-D-glycerate = ADP + 3-

RT Phospho-D-glyceroyl phosphate.

CC -1- PARTWAY: SECOND STEP IN THE SECOND PHASE OF GLYCOLYSIS.

CC -1- SUBUNIT: MONOMER.

CC -1- SUBCELLULAR LOCATION: Cytoplasmic.

CC -1- SIMILARITY: BELONGS TO THE PHOSPHOGLYCERATE KINASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@sib-sib.ch).

DR EMBL: X54519; CAA38375.1; -

DR EMBL: M87647; AAA73203.1; -

DR EMBL: M87648; AAA73206.1; -

DR PIR: S13125; KIBSGM.

DR PIR: J01954; J01954.

DR HSSP: P18912; 1PHF.

DR InterPro: IPR001576; PGK.

DR Pfam: PF00162; PGK.1.

DR PRINTS: PR00477; PHGLYCKINASE.

DR PROSITE: PS00111; GLYCERATE_KINASE.1.

KW transferase; kinase; glycolysis.

SQ SEQUENCE 394 AA; 42457 MW; 7A2E6B978FA7008B CRC64;

Query Match 0.7%; Score 8; DB 1; Length 394;

Best Local Similarity 100.0%; Pred. No. 5.9;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 632 OGAKVILA 639

DB 50 OGAKVILA 57

RESULT 2

NUOG_STRCO STANDARD; PRT: 843 AA.

AC 09XAR0;

DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE NADH dehydrogenase I chain G (EC 1.6.5.3) (NADH-ubiquinone

DE oxidoreductase chain G).

GN NUOG OR SC10A7.15C.

OS Streptomyces coelicolor.

OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;

OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.

OX NCBI_TaxID=1902;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-A3(2);

RA Saunders D.C., Harris D., Bentley S.D., Parkhill J., Barrell B.G.,

RA Rajandream M.A.;

RL Submitted (JUN-1999) to the EMBL/Genbank/DBD databases.

CC -1- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.

CC -1- COFACTOR: BINDS 1 2FE-2S CLUSTER AND 1 4FE-4S CLUSTER (potential).

CC -1- SIMILARITY: BELONGS TO THE COMPLEX I 75 kDa SUBUNIT FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@sib-sib.ch).

DR EMBL: AL076818; CAB44525.1; -

DR InterPro: IPR000283; ComplexI_75k.

DR InterPro: IPR001041; Ferredoxin.

DR InterPro: IPR001467; Molybdopterin.

DR Pfam: PF00384; molybdopterin; 2.

DR PROSITE: PS00641; COMPLEX1_75K.1; 1.

DR PROSITE: PS00642; COMPLEX1_75K.2; 1.

DR PROSITE: PS00643; COMPLEX1_75K.3; 1.

KW Oxidoreductase; NAD; Ubiquinone; Iron-sulfur; 4Fe-4S.

FT METAL 54 54 IRON-SULFUR (2FE-2S) (POTENTIAL).

FT METAL 65 65 IRON-SULFUR (2FE-2S) (POTENTIAL).

FT METAL 68 68 IRON-SULFUR (2FE-2S) (POTENTIAL).

FT METAL 120 120 IRON-SULFUR (2FE-2S) (POTENTIAL).

FT METAL 123 123 IRON-SULFUR (2FE-2S) (POTENTIAL).

FT METAL 129 129 IRON-SULFUR (2FE-2S) (POTENTIAL).

FT METAL 169 169 IRON-SULFUR (4FE-4S) (POTENTIAL).

FT METAL 172 172 IRON-SULFUR (4FE-4S) (POTENTIAL).

FT METAL 175 175 IRON-SULFUR (4FE-4S) (POTENTIAL).

FT METAL 219 219 IRON-SULFUR (4FE-4S) (POTENTIAL).

SQ SEQUENCE 843 AA; 88861 MW; 318D67C2/AAA5193 CRC64;

Query Match 0.7%; Score 8; DB 1; Length 843;
 Best Local Similarity 100.0%; Pred. No. 11;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 104 KDGVLAV 111
 |||||
 DB 775 KDGVLAV 782

RESULT 3

LCK_DROME STANDARD; PRT; 15 AA.
 ID LCK_DROME
 AC P81829;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Leucokinin (DLK).
 GN PP OR DLK.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Neurosecretory cell;
 RX MEDLINE=20044845; Pubmed=10574744;
 RA Terhaz S., O'Connell F.C., Pollock V.P., Kean L., Davies S.A.,
 RA Veenstra J.A., Dow J.A.T.;
 RT "Isolation and characterization of a leucokinin-like peptide of
 RT Drosophila melanogaster."
 RL J. Exp. Biol. 202:3667-3676(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RA Ceuliker S.E., Agbayani A., Arcaina T.T., Baxter E., Blazej R.G.,
 RA Burenhoff C., Champe M., Chavez C., Chew M., Ciesiolka L., Doyle C.M.,
 RA Farfan D.E., Galle R., George R.A., Harris N.L., Hoskins R.A.,
 RA Houston K.A., Hummasti S.R., Karra K., Kearney L., Kim E., Lee B.,
 RA Lewis S., Li P., Lomoten M.A., Mazda P., Moshrefi A.R., Moshrefi M.,
 RA Nixon K., Pacleb J.M., Park S., Pfeiffer B., Poon L., Sequelra A.,
 RA Sehl H., Snir E., Svirskaas R.R., Wan K.H., Weinburg T., Zhang R.,
 RA Zieran L.L., Rubin G.M.;
 RL Submitted (FEB-1999) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: ACTS THROUGH INTRACELLULAR CALCIUM IN MALPIGHIAN TUBULE
 CC -1- STELLATE CELLS TO RAISE CHLORIDE CONDUCTANCE.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: AC006496; -; NOT_ANNOTATED_CDS.
 DR FlyBase: FBgn0028418; Leucokinin.
 DR Neuropeptide; Amidation.
 KW MOD_RES 15 15 AMIDATION.
 FT SEQUENCE 15 AA; 1743 MW; 4793A08F251C9525 CRC64;

Query Match 0.7%; Score 7; DB 1; Length 15;
 Best Local Similarity 100.0%; Pred. No. 3.9;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 243 SVVLGKK 249
 |||||
 DB 2 SVVLGKK 8

RESULT 4
 YACL_ECOLI STANDARD; PRT; 120 AA.
 ID YACL_ECOLI
 AC P45567; P75653;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein yacL.
 GN YACL OR B0119.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / W3110;
 RX MEDLINE=9426130; Pubmed=8202364;
 RA Fujita N., Mori H., Yura T., Ishihama A.;
 RT "Systematic sequencing of the Escherichia coli genome: analysis of
 RT the 2.4-4.1 min (110,917-193,643 bp) region."
 RL Nucleic Acids Res. 22:1637-1639(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RX MEDLINE=97426617; Pubmed=9278503;
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12."
 RL Science 277:1453-1474(1997).
 RN [3]
 RP SEQUENCE OF 18-120 FROM N.A.
 RX MEDLINE=89327165; Pubmed=2666401;
 RA Xie Q.W., Tabor C.W., Tabor H.;
 RT "Spermidine biosynthesis in Escherichia coli: promoter and
 RT termination regions of the speed operon."
 RL J. Bacteriol. 171:4457-4465(1989).
 RN [4]
 RP IDENTIFICATION.
 RX MEDLINE=96032851; Pubmed=7567469;
 RA Bordovsky M., McIninch J., Koonin E.V., Rudd K.E., Medigue C.,
 RA Danchin A.;
 RT "Detection of new genes in a bacterial genome using Markov models for
 RT three gene classes."
 RL Nucleic Acids Res. 23:3554-3562(1995).
 CC -1- SIMILARITY: STRONG, TO H.INFLUENZAE H11724.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: D26562; -; NOT_ANNOTATED_CDS.
 DR EMBL: AE000121; AAC73230.1; ALT_INIT.
 DR EMBL: J02804; -; NOT_ANNOTATED_CDS.
 DR EcoGene; EGI2605; yacL.
 KW Hypothetical protein; Complete proteome.
 FT SEQUENCE 120 AA; 13942 MW; A72F0B3ACDDAE03 CRC64;

Query Match 0.7%; Score 7; DB 1; Length 120;
 Best Local Similarity 100.0%; Pred. No. 24;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 824 QLEFAGD 830
 |||||
 DB 78 QLEFAGD 84

```

RESULT 5
RT17_YEAST STANDARD; PRT: 131 AA.
AC P28778;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Mitochondrial 40S ribosomal protein MRPI7.
GN MRPI7 OR YKL003C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93062809; PubMed=1279374;
RA Haffner P.T., Fox T.D.;
RT "Suppression of carboxy-terminal truncations of the yeast
RT mitochondrial mRNA-specific translational activator PET122 by
RT mutations in two new genes, MRPI7 and PET127."
RL Mol. Gen. Genet. 235:64-73(1992).
RN [2]
RP SEQUENCE FROM N.A.
RA Boyer J., Pascolo S., Richard G.F., Ghazvini M., Colleaux L.,
RA Thierry A., Monnier A.L., Dujon B.;
RL Submitted (MAR-1994) to the EMBL/Genbank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=96242157; PubMed=8668135;
RA Hashida-Okado T., Ogawa A., Endo M., Yasumoto R., Takesako K.,
RA Kato I.;
RT "AUR1, a novel gene conferring aureobasidin resistance on
RT Saccharomyces cerevisiae: a study of defective morphologies in
RT Aur1p-depleted cells."
RL Mol. Genet. 251:236-244(1996).
CC -!- FUNCTION: COMPONENT OF THE SMALL SUBUNIT OF MITOCHONDRIAL
CC RIBOSOME.
CC -!- SUBCELLULAR LOCATION: Mitochondrial.
CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X58362; CAA41256.1; -
DR EMBL: Z28003; CAA81835.1; -
DR EMBL: U49090; AAB06941.1; -
DR PIR: S30119; S30119.
DR SGD: S0001486; MRPI7.
KW Ribosomal protein; Mitochondrion.
SQ SEQUENCE 131 AA; 15021 MW; FEEDB82FEBED7A89F CRC64;

```

```

Query Match 0.7%; Score 7; DB 1; Length 131;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1047 NEDYOSI 1053
Db 125 NEDYOSI 131

```

```

RESULT 6
ID RL15_ECOLI STANDARD; PRT: 144 AA.
AC P02413;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)

```

```

DE 50S ribosomal protein L15.
GN RPLO OR B3301.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=83220807; PubMed=622285;
RA Ceretti D.P., Dean D., Davis G.R., Bedwell D.M., Nomura M.;
RT "The spc ribosomal protein operon of Escherichia coli: sequence and
RT cotranscription of the ribosomal protein genes and a protein export
RT gene."
RL Nucleic Acids Res. 11:2599-2616(1983).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=97426617; PubMed=9278503;
RA Blatner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12."
RL Science 277:1453-1474(1997).
RN [3]
RP SEQUENCE.
RX MEDLINE=78084799; PubMed=340263;
RA Giordano S., Chen R.;
RT "The primary structure of protein L15 located at the
RT peptidyltransferase center of Escherichia coli ribosomes."
RL FEBS Lett. 84:347-350(1977).
RN [4]
RP MASS SPECTROMETRY.
RX MEDLINE=99196679; PubMed=10094780;
RA Arnold R.J., Reilly J.P.;
RT "Observation of Escherichia coli ribosomal proteins and their
RT posttranslational modifications by mass spectrometry."
RL Anal. Biochem. 269:105-112(1999).
CC -!- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 23S RIBOSOMAL RNA.
CC -!- MASS SPECTROMETRY: MM=14980.1; METHOD=MALDI.
CC -!- SIMILARITY: BELONGS TO THE L15P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X01563; CAA25724.1; -
DR EMBL: U18997; AAA58098.1; -
DR EMBL: AE000408; AAC76326.1; -
DR PIR: A02794; R5EC15.
DR Ecogen; EG10876; IPLO.
DR InterPro; IPR001196; Ribosomal_L15.
DR Pfam; PF00256; L15; 1.
DR PROSITE; PS00475; RIBOSOMAL_L15; 1.
KW Ribosomal protein; rRNA-binding; Complete proteome.
SQ SEQUENCE 144 AA; 14980 MW; 53D14CD948B15FD9 CRC64;

```

```

Query Match 0.7%; Score 7; DB 1; Length 144;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 634 AKVILAG 640
Db 108 AKVILAG 114

```

```

RESULT 7
RL5_HAEIN STANDARD: PRT: 144 AA.
AC P44533:
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 50S ribosomal protein L15.
CN RPL0 OR RPL15 OR H10797.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RD / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Felschmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Keriavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodok A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Ulterback T.R., Hanna M.C., Nguyen D.T., Saudke D.M., Bratton R.C.,
RA Fine L.D., Filchman J.L., Fuhrmann J.L., Geoghegan N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus
RT influenzae Rd."
RL Science 269:496-512(1995).
CC -1- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 23S RIBOSOMAL RNA
CC (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE L15P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U32762; AAC22455.1; -
CC DR TIGR: H10797; -
CC DR InterPro: IPR001196; RIBOSOMAL_L15.
CC DR Pfam: PF00256; L15; 1.
CC DR PROSITE: PS00475; RIBOSOMAL_L15; 1.
CC DR RIBOSOMAL protein; rRNA-binding; Complete proteome.
CC KW SEQUENCE 144 AA; 15072 MW; F0CFC80684DC64C5 CRC64;
SQ
Query Match 0.7%; Score 7; DB 1; Length 144;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 634 AKYIILAG 640
DB 108 AKYIILAG 114
RESULT 8
FMO_MORBO STANDARD: PRT: 157 AA.
AC P07640:
DT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Fimbrial protein Q precursor (Beta pilin) (Q pilin).
CN TFPO.
OS Moraxella bovis.
OC Bacteria; Proteobacteria; gamma subdivision; Moraxellaceae; Moraxella.
OX NCBI_TaxID=476;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=EPP63;
RX MEDLINE=85234350; PubMed=2861194;
RA Marrs C.F., Schoolnik G., Koomey J.M., Hardy J., Rothbard J.,
RA Falkow S.;
RT "Cloning and sequencing of a Moraxella bovis pilin gene."
RT J. Bacteriol. 163:132-139(1985).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=EPP63;
RX MEDLINE=90094235; PubMed=2403542;
RA Fuks K.A., Marrs C.F., Stevens S.P., Green M.R.;
RT "Sequence analysis of the inversion region containing the pilin genes
RT of Moraxella bovis."
RT J. Bacteriol. 172:310-316(1990).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=91286182; PubMed=2061282;
RA Rozsa F.W., Marrs C.F.;
RT "Interesting sequence differences between the pilin gene inversion
RT regions of Moraxella lacunata ATCC 17956 and Moraxella bovis Epp63."
RT J. Bacteriol. 173:4000-4006(1991).
RN [4]
RP SEQUENCE OF 7-157.
RX MEDLINE=89010522; PubMed=2902184;
RA Ruelh W.W., Marrs C.F., Fernandez R., Falkow S., Schoolnik G.K.;
RT "Purification, characterization, and pathogenicity of Moraxella bovis
RT pilin."
RL J. Exp. Med. 168:983-1002(1988).
CC -1- SUBUNIT: THE PILI ARE POLAR FLEXIBLE FILAMENTS OF ABOUT 5.4
CC NANOMETERS DIAMETER AND 2.5 MICROMETERS AVERAGE LENGTH; THEY
CC CONSIST OF ONLY A SINGLE POLYPEPTIDE CHAIN ARRANGED IN A HELICAL
CC CONFIGURATION OF FIVE SUBUNITS PER TURN IN THE ASSEMBLED PILUS.
CC -1- MISCELLANEOUS: MORAXELLA BOVIS CAN EXPRESS EITHER A Q OR A I
CC PILIN, THE INVERSION OF 2 KB OF DNA DETERMINES WHICH PILIN IS
CC EXPRESSED.
CC -1- SIMILARITY: BELONGS TO THE N-ME-PHE PILI FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration -
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M11435; AAA25304.1; -
CC DR EMBL: M32345; AAA88223.1; -
CC DR EMBL: M59712; AAA25308.1; -
CC DR PIR: A24434; A24434.
CC DR PIR: J10072; J10072.
CC DR InterPro: IPR001082; Pilin.
CC DR InterPro: IPR001120; Prok_N_methyltn.
CC DR Pfam: PF00114; pilin; 1.
CC DR PRODOM: PD000666; pilin; 1.
CC DR PROSITE: PS00409; PROKAR_NTER_METHYL; 1.
CC KW Fimbria; Methylation.
CC FT PROPEP 1 6
CC FT CHAIN 7 157
CC FT MOD_RES 7 7
CC FT DISULFID 136 155
CC FT METHYLATION.
CC FT BY SIMILARITY.
SQ SEQUENCE 157 AA; 16006 MW; A923CD8A26C693C9 CRC64;
Query Match 0.7%; Score 7; DB 1; Length 157;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 481 ALFEGK 487
DB 55 ALFEGK 61

```

```

RESULT 9
FMI_MORBO STANDARD: PRT: 159 AA.
AC P20657;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Fibritrial protein I precursor (Alpha pilin) (I pilin).
OS Moraxella bovis.
OC Bacteria; Proteobacteria; gamma subdivision; Moraxellaceae; Moraxella.
OX NCBI_TaxID=476;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EBP63;
RX MEDLINE=90094235; PubMed=2403542;
RA Fults K.A., Marrs C.F., Stevens S.P., Green M.R.;
RT "Sequence analysis of the inversion region containing the pilin genes
of Moraxella bovis.";
RL J. Bacteriol. 172:310-316(1990).
RN [2]
RP SEQUENCE OF 7-159.
RX MEDLINE=89010522; PubMed=2902184;
RA Ruehl W.W., Marrs C.F., Fernandez R., Falkow S., Schoolnik G.K.;
RT "Purification, characterization, and pathogenicity of Moraxella bovis
pilin.";
RL J. Exp. Med. 168:983-1002(1988).
CC -1- SUBUNIT: THE PILI ARE POLAR FLEXIBLE FILAMENTS OF ABOUT 5.4
NANOMETERS DIAMETER AND 2.5 MICROMETERS AVERAGE LENGTH; THEY
CONSIST OF ONLY A SINGLE POLYPEPTIDE CHAIN ARRANGED IN A HELICAL
CONFIGURATION OF FIVE SUBUNITS PER TURN IN THE ASSEMBLED PLUS.
CC -1- MISCELLANEOUS: MORAXELLA BOVIS CAN EXPRESS EITHER A O OR A I
PILIN, THE INVERSION OF 2 KB OF DNA DETERMINES WHICH PILIN IS
EXPRESSED.
CC -1- SIMILARITY: BELONGS TO THE N-ME-PHE PILI FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL: M33345; NOT_ANNOTATED_CDS.
CC DR PIR: J10071; J10071.
CC DR HSSP: P02974; IAY2.
CC DR InterPro: IPR001082; Pilin.
CC DR InterPro: IPR001120; Prok_N_methyltn.
CC DR Pfam: PF00114; Pilin; 1.
CC DR ProDom: PD000666; Pilin; 1.
CC DR PROSITE: PS00409; PROKAR_MTER_METHYL; 1.
CC FT FIMBRIA; Methylation.
CC FT PROPEP 1 159 FIMBRIAL PROTEIN I.
CC FT CHAIN 7 159 METHYLATION.
CC FT MOD_RES 7 159 K->KSK (TN REF. 2).
CC FT CONFLICT 159 159 K->KSK (TN REF. 2).
CC SQ SEQUENCE 159 AA; 16723 MW; 9130E2289C7679E CRC64;

Query Match 0.7%; Score 7; DB 1; Length 159;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 481 AALFEGK 487
Db 55 AALFEGK 61

```

```

DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Myosin catalytic light chain LC-1, mantle muscle.
OS Todarodes pacificus (Japanese flying squid).
OC Eukaryota; Metazoa; Mollusca; Cephalopoda; Coleoidea; Teuthoidea.
OC Oegopsida; Ommastrephidae; Todarodes.
OX NCBI_TaxID=6637;
RN [1]
RP SEQUENCE.
RX MEDLINE=87076038; PubMed=3790251;
RA Watanabe B., Maeta T., Kono K., Matsuda G.;
RT "Amino acid sequence of LC-1 light chain of squid mantle muscle
myosin.";
RL Biol. Chem. Hoppe-Seyler 367:1025-1032(1986).
CC -1- FUNCTION: IN MOLLUSCAN MUSCLE, CALCIUM REGULATION IS ASSOCIATED
WITH MYOSIN RATHER THAN WITH ACTIN. MUSCLE MYOSIN CONTAINS TWO
TYPES OF LIGHT CHAINS: THE CATALYTIC LIGHT CHAIN, ESSENTIAL FOR
ATPASE ACTIVITY, AND THE REGULATORY LIGHT CHAIN, A CALCIUM-BINDING
PROTEIN RESPONSIBLE FOR CA(++) DEPENDENT BINDING AND CA(++)
DEPENDENT MG-ATPASE ACTIVITY.
CC -1- SIMILARITY: TO OTHER EF-HAND CALCIUM BINDING PROTEINS, BUT THIS
PROTEIN DOES NOT BIND CALCIUM.
CC DR PIR: A25571; A25571.
CC DR HSSP: P07291; IMDC.
CC DR InterPro: IPR002048; EF-hand.
CC DR Pfam: PF00036; ethand; 1.
CC KW Myosin; Muscle protein.
CC FT MOD_RES 1 1 BLOCKED.
CC SQ SEQUENCE 159 AA; 18059 MW; 0F60334A34971410 CRC64;

Query Match 0.7%; Score 7; DB 1; Length 159;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 923 AKVGDL 929
Db 31 AKVGDL 37

RESULT 11
POPS_YEAST STANDARD: PRT: 173 AA.
AC P28005;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ribonucleases P/MRP protein subunit POPS (EC 3.1.26.5) (RNases P/MRP
DE 19.6 kDa subunit) (RNA processing protein POPS).
GN POPS OR YAL033W OR FUN53.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92260538; PubMed=1583694;
RA Harris S.D., Cheng J., Pugh T.A., Pringle J.R.;
RT "Molecular analysis of Saccharomyces cerevisiae chromosome I. On the
number of genes and the identification of essential genes using
temperature-sensitive-lethal mutations.";
RL J. Mol. Biol. 225:35-65(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=5288C / AB972;
RX MEDLINE=95249563; PubMed=7731988;
RA Bussey H., Kaback D.B., Zhong W., Vo D.T., Clark M.W., Fortin N.,
RA Hall J., Ouellette B.F.F., Keng T., Barton A.B., Su Y., Davies C.K.,
RA Storms R.K.;
RT "The nucleotide sequence of chromosome I from Saccharomyces
cerevisiae.";
RL Proc. Natl. Acad. Sci. U.S.A. 92:3809-3813(1995).
CC -1- FUNCTION: COMPONENT OF RIBONUCLEASE P, A PROTEIN COMPLEX THAT

```



```

CC GENERATES MATURE tRNA MOLECULES BY CLEAVING THEIR 5' ENDS.
CC ALSO A COMPONENT OF RNASE MRP.
CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage of RNA, removing 5'-
CC extra-nucleotide from tRNA precursor.
CC -1- SUBUNIT: COMPONENT OF NUCLEAR RNASE P AND RNASE MRP RNASE P
CC RIBONUCLEOPROTEINS. RNASE P CONSISTS OF A RNA MOIETY AND AT LEAST
CC 8 PROTEIN SUBUNITS; POP1, POP3, POP4, POP5, POP6, POP7, POP8 AND
CC RPP1.
CC -1- SUBCELLULAR LOCATION: Nuclear (potential).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL; X62577; CAA44457.1; -.
CC DR EMBL; U12980; AAC0499.1; -.
CC DR PIR; S23411; S23411.
CC DR SGD; S0000031; POP5.
CC DR InterPro; IPR002759; DUF69.
CC DR Pfam; PF01900; DUF69; 1.
CC DR PRODOM; PD012772; DUF69; 1.
CC KW Hydrolase; Nuclear protein; tRNA processing.
CC SEQUENCE 173 AA; 19573 MW; 918193631BD790DD CRC64;

Query Match 0.7%; Score 7; DB 1; Length 173;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1004 LGDYGSA 1010
DB 60 LGDYGSA 66
|||||

```

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL; X55387; CAA39058.1; -.
CC DR PIR; S12256; RDOFBR.
CC DR HSSP; P13272; IRIE.
CC DR InterPro; IPR01281; Rieske.
CC DR Pfam; PF00355; Rieske; 1.
CC DR PRINTS; PR00162; Rieske.
CC DR PROSITE; PS00199; RIESKE_1; 1.
CC DR PROSITE; PS00200; RIESKE_2; 1.
CC KW Electron transport; Inner membrane; Transmembrane; Iron-sulfur;
CC Oxidoreductase.
CC FT INIT_MET 0
CC FT TRANSMEM 20 40
CC FT METAL 120 120 POTENTIAL.
CC FT METAL 122 122 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
CC FT METAL 144 144 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
CC FT METAL 147 147 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
CC FT DISULFID 125 146 BY SIMILARITY.
CC SEQUENCE 182 AA; 19382 MW; 4AC1F24D74CE3A4 CRC64;

Query Match 0.7%; Score 7; DB 1; Length 182;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 526 DRLVYGT 532
DB 18 DRLVYGT 24
|||||

```

```

RESULT 12
UCRI_RHORU STANDARD: PRT; 182 AA.
AC P23136;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 15-JUL-1989 (Rel. 38, Last annotation update)
DE Ubiquinol-cytochrome c reductase iron-sulfur subunit (EC 1.10.2.2)
DE (Rieske iron-sulfur protein) (RISP).
GN PETA OR FBCE.
OS Rhodospirillum rubrum.
OC Bacteria; Proteobacteria; alpha subdivision; Rhodospirillaceae;
OC Rhodospirillum.
OX NCBI_TaxID=1085;
RN [1]
RN SEQUENCE FROM N.A., AND SEQUENCE OF 1-34.
RC STRAIN=FR1;
RX MEDLINE=91094774; PubMed=2176269;
RA Majewski C., Trebst A.;
RT "The pet genes of Rhodospirillum rubrum: cloning and sequencing of
RT the genes for the cytochrome bcl-complex.";
RL Mol. Gen. Genet. 224:373-382(1990).
CC -1- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE
CC COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A
CC RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL
CC COUPLED TO ATP SYNTHESIS.
CC -1- CATALYTIC ACTIVITY: QH(2) + 2 ferricytochrome c = Q + 2
CC ferricytochrome c.
CC -1- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B,
CC CYTOCHROME C1 AND THE RIESKE PROTEIN.
CC -1- SUBCELLULAR LOCATION: Membrane-bound.
CC -1- MISCELLANEOUS: THE RIESKE PROTEIN IS A HIGH POTENTIAL 2FE-2S
CC PROTEIN.
CC -1- SIMILARITY: TO RIESKE PROTEINS FROM OTHER SOURCES (MITOCHONDRIA,
CC BACTERIAL, CHLOROPLAST).

```

```

RESULT 13
YEIP_ECOLI STANDARD: PRT; 190 AA.
ID YEIP_ECOLI
AC P33028;
DT 01-FEB-1994 (Rel. 28, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Protein yeip.
DE Protein yeip.
GN YEIP OR B2171.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=K12 / BHB2600;
RA Richterich P., Lakey N., Gryan G., Jaehn L., Mintz L., Robison K.,
RA Church G.M.;
RL Submitted (OCT-1993) to the EMBL/GenBank/DBJ databases.
RN [2]
RN SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [3]
RN SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=97251358; PubMed=9097040;
RA Itoh T., Alba H., Baba T., Fujita K., Hayashi K., Inada T.,
RA Isono K., Kasai H., Kimura S., Kitakawa M., Kitagawa M.,

```

```

RA Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,
RA Nakabe S., Nakamura Y., Nishimoto H., Nishio Y., Oshima T.,
RA Salto N., Samped G., Seki Y., Sivasubraman S., Tagami H.,
RA Takeda J., Takemoto K., Wada C., Yamamoto Y., Horikuchi T.;
RT "A 460-kb DNA sequence of the Escherichia coli K-12 genome
RT corresponding to the 40.1-50.0 min region on the linkage map.";
RL DNA Res. 3:379-392(1996).
RN [4]
RP SEQUENCE OF 50-98 FROM N.A.
RC STRAIN-K12 / W3110;
RX MEDLINE=94162733; PubMed=7764507;
RA Yamada M., Yanai S., Talkuder A.;
RT "Analysis of products of the Escherichia coli genomic genes and
RT regulation of their expressions: an applicable procedure for genomic
RT analysis of other microorganisms.";
RL Biosci. Biotechnol. Blochem. 58:117-120(1994).
RN [5]
RP IDENTIFICATION BY MASS SPECTROMETRY.
RX MEDLINE=99420866; PubMed=10493123;
RA Fountoulakis M., Takacs M.-F., Berndt P., Langen H., Takacs B.;
RT "Enrichment of low abundance proteins of Escherichia coli by
RT hydroxyapatite chromatography.";
RL Electrophoresis 20:2181-2195(1999).
CC -I- SIMILARITY: BELONGS TO THE ELONGATION FACTOR P FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; U00007; AAA60519.1; ALT_INIT.
DR EMBL; AE000306; AAC75232.1; ALT_INIT.
DR EMBL; D90849; BAA15980.1; ALT_INIT.
DR EMBL; D21148; BAA04684.1; -.
DR Ecocore: EG12035; YelP.
DR InterPro: IPR001059; EFP.
DR Pfam; PF01132; EFP; 1.
DR PROSITE; PS01275; EFP; 1.
DR Complete proteome.
KW SEQUENCE 190 AA; 21532 MW; 628E9E8A26EC7CA CRC64;
SQ
Query Match 0.7%; Score 7; DB 1; Length 190;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 370 DGQLLAL 376
DB 121 DGQLLAL 127
ID SODF_TETPY STANDARD: PRT; 196 AA.
AC P19666;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last annotation update)
DE Superoxide dismutase [Fe] (EC 1.15.1.1).
OS Tetrahymena pyriformis.
OC Eukaryota; Alveolata; Ciliophora; Oligohymenophorea; Hymenostomatida;
OC Tetrahymena; Tetrahymena.
OX NCBI_TaxID=5908;
RN [1]
RP SEQUENCE.
RX MEDLINE=91009226; PubMed=2170391;
RA Barra D., Schinina M.E., Bossa F., Puget K., Durosay P., Guissani A.,
RA Michelson A.M.;
RT "A tetrameric iron superoxide dismutase from the eucaryote
RT Tetrahymena pyriformis.";
```

Query Match	Best Local	Similarity	0.7%	Score 7	DB 1	Length 196			
Matches	7	Conservative	0	Mismatches	0	Indels	0	Gaps	0
QY	1060	KKRYLQA	1066						
Db	188	KKRYLQA	194						
RESULT	15								
DEFRA_MEDSA	STANDARD	PRT	217 AA						
AC	P51109								
DT	01-OCT-1996	(Rel. 34, Created)							
DT	01-OCT-1996	(Rel. 34, Last sequence update)							
DT	01-OCT-1996	(Rel. 34, Last annotation update)							
DE	Dihydroflavonol 4-reductase (EC 1.1.1.219) (DFR) (Dihydrokaempferol 4-reductase) (Fragment).								
GN	DFRL								
OS	Medicago sativa (Alfalfa).								
CC	Euarvota, Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eucosids I; Fabales; Fabaceae; Papilionoideae; Trifoliaceae; Medicago.								
OX	NCBI_TaxID=3879;								
RP	SEQUENCE FROM N.A.								
RX	STRAIN=CV, VARIA A2;								
RA	MEDLINE=96128019; PubMed=8541503;								
RT	Charrier B., Coronado C., Kondorosi A., Ratet P.; "Molecular characterization and expression of alfalfa (Medicago sativa L.) flavanone-3-hydroxylase and dihydroflavonol 4-reductase encoding genes."								
RL	Plant Mol. Biol. 29: 773-786(1995).								
CC	-1- CATALYTIC ACTIVITY: Cis-3,4-Leucopelargonidin + NADP(+) = (+)-dihydrokaempferol + NADPH.								
CC	-1- SIMILARITY: BELONGS TO THE DIHYDROFLAVONOL-4-REDUCTASES FAMILY.								
CC	-1- PATHWAY: FLAVONOID SYNTHESIS; ANTHOCYANIDINS BIOSYNTHESIS.								
CC	-1- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE FAMILY.								
CC	PIR: A39223; A39223.								
DR	HSSP; P04179; IAP6.								
DR	InterPro: IPR001189; SOD_M1.								
DR	Pfam: PF00081; sodfe_1.								
DR	Pfam: PF02777; sodfe_C_1.								
DR	ProDom: PD000475; SOD_M1; 1.								
DR	PROSITE: PS00088; SOD_MN; 1.								
RW	Oxidoreductase; Iron.								
FT	METAL	20							
FT	METAL	68							
FT	METAL	157							
FT	METAL	161							
SO	SEQUENCE	196 AA; 22657 MW; FCIFF2P67893DBDC7 CRC64;							

SO SEQUENCE 217 AA: 24389 MM: EBA43AA76CBEE3EB2 CRC64:

Query Match 0.7%; Score 7; DB 1; Length 217;

Best Local Similarity 100.0%; Pred. No. 40%;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 596 PTIKGYL 602
|||||

Db 85 PTIKGYL 91

Search completed: May 21, 2002, 14:55:21
Job time: 180 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 21, 2002, 14:53:21 ; Search time 42.6 Seconds
(without alignments)
4361.421 Million cell updates/sec

Title: US-09-729-653-2_COPY_1_1074
Perfect score: 1074
Sequence: 1 HSLIGRCSRLLDGNVAVC.....LYFEGEKRYLQAKFFLLCG 1074

Scoring table:
Gapop 60.0 , Gapext 60.0

Searched: 562222 seqs, 172994929 residues

Word size : 0

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :

SPTREMBL_19:*
1: sp-archaea:*
2: sp-bacteria:*
3: sp-fungi:*
4: sp-human:*
5: sp-invertebrate:*
6: sp-mammal:*
7: sp-mhc:*
8: sp-organelle:*
9: sp-phage:*
10: sp-plant:*
11: sp-rodent:*
12: sp-virus:*
13: sp-vertebrate:*
14: sp-unclassified:*
15: sp-virus:*
16: sp-bacteriap:*
17: sp-archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	498	46.4	905	4	Q9HCD4
2	270	25.1	481	4	Q9H5S0
3	9	0.8	519	5	Q46018
4	9	0.8	706	5	Q9GY76
5	8	0.7	120	15	Q95SR1
6	8	0.7	176	12	Q84636
7	8	0.7	241	17	Q9H087
8	8	0.7	279	16	Q98FH3
9	8	0.7	300	2	Q93F73
10	8	0.7	300	2	Q93F70
11	8	0.7	474	13	Q73906
12	8	0.7	493	2	Q9EV50
13	8	0.7	505	2	Q9KM19
14	8	0.7	519	16	Q9CD12
15	8	0.7	699	12	Q993G4
16	8	0.7	1214	10	Q9LD60

17	8	0.7	1228	5	Q61198	Q61198 caenorhabdi
18	8	0.7	1716	11	Q99M57	Q99M57 mus muscula
19	8	0.7	1750	5	Q9BK11	Q9BK11 trypanosoma
20	7	0.7	39	16	Q927N5	Q927N5 rhizobium m
21	7	0.7	85	15	Q9YUK0	Q9YUK0 human immun
22	7	0.7	94	10	Q9SK19	Q9SK19 arabidopsis
23	7	0.7	98	8	Q04290	Q04290 phaseolus v
24	7	0.7	98	8	Q9MEV2	Q9MEV2 sclerous vul
25	7	0.7	100	3	Q96U24	Q96U24 neurospora
26	7	0.7	105	9	Q9B061	Q9B061 mycobacteri
27	7	0.7	110	16	Q930C5	Q930C5 rhizobium m
28	7	0.7	118	3	Q9U022	Q9U022 sorcharia fi
29	7	0.7	118	3	Q9U0U9	Q9U0U9 sorcharia ma
30	7	0.7	118	3	Q9U0U8	Q9U0U8 sorcharia sc
31	7	0.7	118	3	Q9U0U4	Q9U0U4 neurospora
32	7	0.7	118	3	Q9U0X4	Q9U0X4 neurospora
33	7	0.7	118	3	Q9U0X2	Q9U0X2 neurospora
34	7	0.7	118	3	Q9U0W8	Q9U0W8 neurospora
35	7	0.7	118	3	Q9U0W6	Q9U0W6 neurospora
36	7	0.7	118	3	Q9U0W1	Q9U0W1 neurospora
37	7	0.7	118	3	Q9U0W0	Q9U0W0 neurospora
38	7	0.7	118	3	Q9U0X0	Q9U0X0 neurospora
39	7	0.7	118	3	Q9U0W3	Q9U0W3 neurospora
40	7	0.7	118	3	Q9U0X6	Q9U0X6 neurospora
41	7	0.7	122	2	Q86438	Q86438 pseudomonas
42	7	0.7	123	10	Q9L0V5	Q9L0V5 arabidopsis
43	7	0.7	126	3	Q96U28	Q96U28 gelastinospo
44	7	0.7	126	3	Q96U27	Q96U27 gelastinospo
45	7	0.7	126	3	Q96U26	Q96U26 gelastinospo
46	7	0.7	126	3	Q96U25	Q96U25 gelastinospo
47	7	0.7	127	10	Q9FTY2	Q9FTY2 oryza sativ
48	7	0.7	140	10	Q940J7	Q940J7 arabidopsis
49	7	0.7	140	16	Q9CKR0	Q9CKR0 pasteurella
50	7	0.7	141	5	Q9VPP7	Q9VPP7 drosophila
51	7	0.7	142	5	Q94753	Q94753 schistosoma
52	7	0.7	143	16	Q9CBJ7	Q9CBJ7 mycobacteri
53	7	0.7	143	3	Q14368	Q14368 schizosacch
54	7	0.7	150	2	Q9Z120	Q9Z120 burkholderi
55	7	0.7	151	12	P88893	P88893 jamaican to
56	7	0.7	153	3	Q9HDD7	Q9HDD7 paracoccidi
57	7	0.7	153	3	Q9HGX6	Q9HGX6 paracoccidi
58	7	0.7	153	5	Q9YUFA	Q9YUFA drosophila
59	7	0.7	154	16	Q9CHW9	Q9CHW9 lactococcus
60	7	0.7	156	2	Q59507	Q59507 moraxella b
61	7	0.7	157	2	Q9F747	Q9F747 bacteroides
62	7	0.7	157	2	Q59503	Q59503 moraxella b
63	7	0.7	158	2	Q59501	Q59501 moraxella b
64	7	0.7	158	2	Q59508	Q59508 moraxella b
65	7	0.7	159	2	Q60163	Q60163 moraxella b
66	7	0.7	159	2	Q59505	Q59505 moraxella b
67	7	0.7	160	2	Q59504	Q59504 moraxella b
68	7	0.7	173	2	Q9F4J9	Q9F4J9 streptococc
69	7	0.7	179	2	Q69605	Q69605 methylobact
70	7	0.7	184	8	Q63042	Q63042 bralithwaite
71	7	0.7	186	10	Q9M737	Q9M737 phyllodrepa
72	7	0.7	188	8	Q63041	Q63041 thuidium ta
73	7	0.7	188	10	Q9M790	Q9M790 trachyloma
74	7	0.7	188	10	Q9M730	Q9M730 bescherelli
75	7	0.7	189	8	Q9BC49	Q9BC49 andreaebry
76	7	0.7	189	8	Q9BC35	Q9BC35 oedipodium
77	7	0.7	189	8	Q9BC27	Q9BC27 schlottheimi
78	7	0.7	189	8	Q9SE40	Q9SE40 tayloria ii
79	7	0.7	189	10	Q9M7A5	Q9M7A5 forstroemi
80	7	0.7	190	2	P96098	P96098 thiodacilli
81	7	0.7	190	10	Q9M772	Q9M772 thuidium de
82	7	0.7	191	8	Q9MUF4	Q9MUF4 dawsonia pa
83	7	0.7	191	8	Q9GEA2	Q9GEA2 campylopus
84	7	0.7	191	8	Q9GEA1	Q9GEA1 campylopus
85	7	0.7	191	8	Q9GE97	Q9GE97 dicranodent
86	7	0.7	191	8	Q9GCE9	Q9GCE9 bescherelli
87	7	0.7	193	10	Q9AY09	Q9AY09 chamaedryum
88	7	0.7	194	8	Q63029	Q63029 tayloria ii
89	7	0.7	201	16	Q50930	Q50930 borrelia bu

90	7	0.7	201	16	Q92527	Q92527 r transposa	163	7	0.7	361	12	Q9YB52	Q9YB52 vigna mungo
91	7	0.7	207	12	Q12668	Q12668 tomato dwar	164	7	0.7	361	12	Q88549	Q88549 tomato leaf
92	7	0.7	214	11	Q63177	Q63177 ratus mus norv	165	7	0.7	361	12	Q91272	Q91272 tomato leaf
93	7	0.7	214	11	Q922H8	Q922H8 mus musculi	166	7	0.7	362	12	Q99DB7	Q99DB7 vigna mungo
94	7	0.7	214	12	Q991R6	Q991R6 tomato leaf	167	7	0.7	362	12	Q91SC1	Q91SC1 indian mung
95	7	0.7	214	12	Q991R4	Q991R4 tomato leaf	168	7	0.7	362	12	Q91P84	Q91P84 tomato leaf
96	7	0.7	219	16	Q921G3	Q921G3 rhizobium m	169	7	0.7	362	12	Q91N70	Q91N70 mungbean ye
97	7	0.7	223	16	Q10843	Q10843 mycobacteri	170	7	0.7	362	12	Q91N71	Q91N71 indian mung
98	7	0.7	226	12	Q9WP12	Q9WP12 havana toma	171	7	0.7	362	12	Q913E9	Q913E9 indian mung
99	7	0.7	228	4	Q9NMG3	Q9NMG3 homo sapien	172	7	0.7	362	12	Q910Y7	Q910Y7 soybean yel
100	7	0.7	229	16	Q9KSR0	Q9KSR0 vibrio chol	173	7	0.7	364	12	Q9W7F8	Q9W7F8 newcastle d
101	7	0.7	232	10	Q22922	Q22922 arabidopsis	174	7	0.7	364	12	Q88VP3	Q88VP3 newcastle d
102	7	0.7	239	2	Q9F9K7	Q9F9K7 pisciticket	175	7	0.7	364	12	Q83B32	Q83B32 newcastle d
103	7	0.7	241	16	Q989J1	Q989J1 rhizobium 1	176	7	0.7	364	12	Q83B33	Q83B33 newcastle d
104	7	0.7	248	17	Q9V069	Q9V069 pyrococcus	177	7	0.7	364	12	Q83B35	Q83B35 newcastle d
105	7	0.7	250	16	Q9KOE5	Q9KOE5 vibrio chol	178	7	0.7	364	12	Q83B38	Q83B38 newcastle d
106	7	0.7	251	2	Q9FDR9	Q9FDR9 microcystis	179	7	0.7	364	12	Q83B39	Q83B39 newcastle d
107	7	0.7	251	2	Q9RNB4	Q9RNB4 microcystis	180	7	0.7	364	12	Q83B40	Q83B40 newcastle d
108	7	0.7	252	10	Q9F120	Q9F120 arabidopsis	181	7	0.7	364	12	Q9WBL3	Q9WBL3 newcastle d
109	7	0.7	256	17	Q97CM5	Q97CM5 thermoplasm	182	7	0.7	364	12	Q90338	Q90338 newcastle d
110	7	0.7	265	1	Q39598	Q39598 chlamydomon	183	7	0.7	364	12	Q92330	Q92330 newcastle d
111	7	0.7	266	2	Q9AKE2	Q9AKE2 rickettsia	184	7	0.7	364	12	Q9WG22	Q9WG22 newcastle d
112	7	0.7	266	2	Q9AKO1	Q9AKO1 rickettsia	185	7	0.7	364	12	Q9WG23	Q9WG23 newcastle d
113	7	0.7	274	5	Q94849	Q94849 drosophila	186	7	0.7	364	12	Q9WG24	Q9WG24 newcastle d
114	7	0.7	277	5	Q94969	Q94969 drosophila	187	7	0.7	364	12	Q9WG25	Q9WG25 newcastle d
115	7	0.7	282	17	Q9UWY6	Q9UWY6 sulfolobus	188	7	0.7	364	12	Q9WG26	Q9WG26 newcastle d
116	7	0.7	285	2	Q9RL43	Q9RL43 streptomyc	189	7	0.7	364	12	Q9WG27	Q9WG27 newcastle d
117	7	0.7	292	2	Q9ACY3	Q9ACY3 streptomyc	190	7	0.7	364	12	Q9WG28	Q9WG28 newcastle d
118	7	0.7	293	10	Q9M053	Q9M053 arabidopsis	191	7	0.7	364	12	Q9WG29	Q9WG29 newcastle d
119	7	0.7	293	10	Q93V35	Q93V35 oryza sativ	192	7	0.7	364	12	Q9WH00	Q9WH00 newcastle d
120	7	0.7	295	2	Q44588	Q44588 alcaligenes	193	7	0.7	364	12	Q9WH01	Q9WH01 newcastle d
121	7	0.7	299	2	Q93F63	Q93F63 bacillus sp	194	7	0.7	364	12	Q9WH02	Q9WH02 newcastle d
122	7	0.7	305	4	Q15287	Q15287 homo sapien	195	7	0.7	364	12	Q9WH03	Q9WH03 newcastle d
123	7	0.7	305	4	Q75308	Q75308 homo sapien	196	7	0.7	364	12	Q9DLD5	Q9DLD5 newcastle d
124	7	0.7	305	11	Q99M28	Q99M28 mus musculi	197	7	0.7	364	12	Q91RF2	Q91RF2 newcastle d
125	7	0.7	306	12	Q91ET9	Q91ET9 cotton leaf	198	7	0.7	364	12	Q914X4	Q914X4 newcastle d
126	7	0.7	310	2	Q9F1K3	Q9F1K3 citrobacter	199	7	0.7	365	17	Q27876	Q27876 methanother
127	7	0.7	310	2	Q9F1K3	Q9F1K3 citrobacter	200	7	0.7	367	4	Q96NB9	Q96NB9 homo sapien
128	7	0.7	310	16	Q9JOL6	Q9JOL6 neisseria m	201	7	0.7	371	16	Q9ZDP9	Q9ZDP9 rickettsia
129	7	0.7	312	16	Q67040	Q67040 aquifex aeo	202	7	0.7	373	4	Q9H6B4	Q9H6B4 homo sapien
130	7	0.7	314	2	Q85929	Q85929 sphingomona	203	7	0.7	373	11	Q920S5	Q920S5 mus musculi
131	7	0.7	314	16	Q9ZJ38	Q9ZJ38 helicobacte	204	7	0.7	379	16	Q05775	Q05775 mycobacteri
132	7	0.7	315	2	Q9R9L5	Q9R9L5 rhizobium m	205	7	0.7	388	5	Q21174	Q21174 caenorhabdi
133	7	0.7	315	2	Q9X983	Q9X983 rhizobium m	206	7	0.7	390	10	Q93Z35	Q93Z35 arabidopsis
134	7	0.7	315	2	Q52887	Q52887 rhizobium m	207	7	0.7	391	4	Q75854	Q75854 homo sapien
135	7	0.7	315	16	Q26091	Q26091 helicobacte	208	7	0.7	392	17	Q58317	Q58317 pyrococcus
136	7	0.7	316	16	Q986T0	Q986T0 rhizobium 1	209	7	0.7	394	16	Q9X213	Q9X213 thermotoga
137	7	0.7	318	4	Q9P0R9	Q9P0R9 homo sapien	210	7	0.7	396	17	Q97ZL6	Q97ZL6 sulfolobus
138	7	0.7	319	4	Q9P0P6	Q9P0P6 homo sapien	211	7	0.7	406	2	Q9X8X9	Q9X8X9 streptomyc
139	7	0.7	322	5	Q9VEH1	Q9VEH1 drosophila	212	7	0.7	409	2	Q8KJG1	Q8KJG1 streptococ
140	7	0.7	330	5	Q16721	Q16721 caenorhabdi	213	7	0.7	412	3	Q9HDF1	Q9HDF1 paracoccidi
141	7	0.7	337	10	Q9GSD9	Q9GSD9 strongyloce	214	7	0.7	413	3	Q9HES0	Q9HES0 paracoccidi
142	7	0.7	337	10	Q9XG67	Q9XG67 nicotiana t	215	7	0.7	413	8	Q9B6C9	Q9B6C9 yarrowia 11
143	7	0.7	338	2	Q9RD96	Q9RD96 streptomyc	216	7	0.7	415	3	Q60060	Q60060 schizosacch
144	7	0.7	338	5	P91027	P91027 caenorhabdi	217	7	0.7	416	3	Q9H0N9	Q9H0N9 paracoccidi
145	7	0.7	340	16	Q9ZXE1	Q9ZXE1 rhizobium m	218	7	0.7	416	3	Q9HDI9	Q9HDI9 paracoccidi
146	7	0.7	344	3	P78613	P78613 emericella	219	7	0.7	416	3	Q9HES3	Q9HES3 paracoccidi
147	7	0.7	344	16	P71057	P71057 bacillus su	220	7	0.7	416	3	Q9HES2	Q9HES2 paracoccidi
148	7	0.7	345	10	Q9SRN7	Q9SRN7 arabidopsis	221	7	0.7	416	3	Q9HES1	Q9HES1 paracoccidi
149	7	0.7	347	3	Q9P956	Q9P956 penicillium	222	7	0.7	416	3	Q9HER9	Q9HER9 paracoccidi
150	7	0.7	348	3	P78579	P78579 emericella	223	7	0.7	416	3	Q9HER8	Q9HER8 paracoccidi
151	7	0.7	350	2	Q43983	Q43983 acetobact	224	7	0.7	416	3	Q9HER7	Q9HER7 paracoccidi
152	7	0.7	352	5	Q9XV05	Q9XV05 caenorhabdi	225	7	0.7	416	3	Q9HER5	Q9HER5 paracoccidi
153	7	0.7	353	2	Q32433	Q32433 acetobact	226	7	0.7	416	3	Q9HER0	Q9HER0 paracoccidi
154	7	0.7	353	2	Q52166	Q52166 pseudomonas	227	7	0.7	425	5	Q9NPD2	Q9NPD2 drosophila
155	7	0.7	353	2	Q52175	Q52175 pseudomonas	228	7	0.7	425	16	Q05301	Q05301 mycobacteri
156	7	0.7	353	3	P87210	P87210 neurospora	229	7	0.7	427	2	Q9EZB0	Q9EZB0 bruceella ab
157	7	0.7	356	16	Q99SCH4	Q99SCH4 staphylococ	230	7	0.7	428	10	Q9ZPF6	Q9ZPF6 arabidopsis
158	7	0.7	358	2	Q93ED3	Q93ED3 rhizobium 1	231	7	0.7	431	2	Q935K0	Q935K0 salmonella
159	7	0.7	359	12	Q67620	Q67620 tomato yell	232	7	0.7	433	10	Q9ZVD8	Q9ZVD8 arabidopsis
160	7	0.7	361	12	Q88949	Q88949 tomato yell	233	7	0.7	435	4	Q9H8G7	Q9H8G7 homo sapien
161	7	0.7	361	12	Q9YRC0	Q9YRC0 angled luff	234	7	0.7	435	5	Q27041	Q27041 thelleria p
162	7	0.7	361	12	Q9WH24	Q9WH24 indian mung	235	7	0.7	435	17	Q9YEB1	Q9YEB1 aeropyrum p

236	7	0.7	436	10	Q9FWH8	Q9FWH8 oryza sativ	309	7	0.7	752	12	P88916	P88916 kaposi's sa
237	7	0.7	439	2	Q9LS16	Q9LS16 salmoneilla	310	7	0.7	752	12	Q40924	Q40924 kaposi's sa
238	7	0.7	439	10	Q9A24	Q9A24 arbidopsis	311	7	0.7	753	5	Q9VL20	Q9VL20 drosophila
239	7	0.7	440	11	Q9R249	Q9R249 mus musculu	312	7	0.7	758	5	Q22203	Q22203 caenorhabd
240	7	0.7	440	16	Q83652	Q83652 treponema p	313	7	0.7	760	2	Q9FC45	Q9FC45 streptomyce
241	7	0.7	441	12	Q994D7	Q994D7 porcine ade	314	7	0.7	768	10	Q9LH00	Q9LH00 oryza sativ
242	7	0.7	446	11	Q99ME7	Q99ME7 mus musculu	315	7	0.7	769	5	Q23275	Q23275 arbidopsis
243	7	0.7	450	4	Q9BY54	Q9BY54 homo sapien	316	7	0.7	777	10	Q9W4K8	Q9W4K8 drosophila
244	7	0.7	450	11	Q9CWU1	Q9CWU1 mus musculu	317	7	0.7	784	10	Q9SJ00	Q9SJ00 arbidopsis
245	7	0.7	460	10	Q23014	Q23014 arbidopsis	318	7	0.7	798	5	Q9VLY9	Q9VLY9 drosophila
246	7	0.7	460	16	P77793	P77793 escherichia	319	7	0.7	803	10	Q23540	Q23540 arbidopsis
247	7	0.7	462	10	Q43742	Q43742 brassica na	320	7	0.7	806	10	Q81610	Q81610 pisum sativ
248	7	0.7	462	10	Q9S7M3	Q9S7M3 arbidopsis	321	7	0.7	806	10	Q9T0M9	Q9T0M9 pisum sativ
249	7	0.7	463	16	Q983M2	Q983M2 rhizobium l	322	7	0.7	807	3	Q94275	Q94275 schizosacch
250	7	0.7	464	8	Q9BRC2	Q9BRC2 amphidinium	323	7	0.7	818	5	Q9YJN0	Q9YJN0 drosophila
251	7	0.7	473	16	Q9PNC5	Q9PNC5 campylobact	324	7	0.7	825	16	Q9PQ14	Q9PQ14 ureaplasma
252	7	0.7	478	4	Q9BSJ7	Q9BSJ7 homo sapien	325	7	0.7	843	2	P72211	P72211 proteus mir
253	7	0.7	480	5	Q9VMD8	Q9VMD8 drosophila	326	7	0.7	846	2	Q9ZFP7	Q9ZFP7 vibrio chol
254	7	0.7	481	16	Q07787	Q07787 mycobacteri	327	7	0.7	846	16	Q9KOP6	Q9KOP6 vibrio chol
255	7	0.7	488	16	Q97J20	Q97J20 clostridium	328	7	0.7	858	10	Q9C7R5	Q9C7R5 arbidopsis
256	7	0.7	489	5	Q22124	Q22124 caenorhabd	329	7	0.7	858	10	Q9C7R4	Q9C7R4 arbidopsis
257	7	0.7	509	16	Q98C99	Q98C99 rhizobium l	330	7	0.7	886	4	Q90MG8	Q90MG8 homo sapien
258	7	0.7	510	4	Q9H1L6	Q9H1L6 homo sapien	331	7	0.7	907	10	Q9SE90	Q9SE90 brassica ol
259	7	0.7	510	16	Q9A748	Q9A748 caulobacter	332	7	0.7	943	17	Q30320	Q30320 archaeoglob
260	7	0.7	511	2	P94786	P94786 ehrlichia r	333	7	0.7	986	2	Q59506	Q59506 microspor
261	7	0.7	511	2	Q9NPL7	Q9NPL7 homo sapien	334	7	0.7	988	12	Q9LEK5	Q9LEK5 fowlpox vir
262	7	0.7	512	4	P94787	P94787 ehrlichia r	335	7	0.7	1042	10	Q9LIM7	Q9LIM7 zea mays (m
263	7	0.7	512	2	Q936C7	Q936C7 streptomyce	336	7	0.7	1055	10	Q9LIV9	Q9LIV9 oryza sativ
264	7	0.7	513	16	Q07415	Q07415 mycobacteri	337	7	0.7	1060	10	Q04709	Q04709 zea mays (m
265	7	0.7	517	5	Q18655	Q18655 caenorhabd	338	7	0.7	1062	5	Q76657	Q76657 caenorhabd
266	7	0.7	520	16	Q99S82	Q99S82 staphylococ	339	7	0.7	1062	5	Q960C4	Q960C4 drosophila
267	7	0.7	520	16	Q931G7	Q931G7 staphylococ	340	7	0.7	1087	5	Q9BL04	Q9BL04 caenorhabd
268	7	0.7	522	16	Q9KDI0	Q9KDI0 bacillus ha	341	7	0.7	1099	2	Q54377	Q54377 lactococcus
269	7	0.7	525	4	Q9H9Z3	Q9H9Z3 homo sapien	342	7	0.7	1151	4	Q9UN15	Q9UN15 homo sapien
270	7	0.7	525	4	Q96176	Q96176 homo sapien	343	7	0.7	1151	11	Q9Z0F2	Q9Z0F2 mus musculu
271	7	0.7	531	10	Q65536	Q65536 arbidopsis	344	7	0.7	1250	5	Q9N595	Q9N595 caenorhabd
272	7	0.7	545	16	Q9PE05	Q9PE05 xyella fas	345	7	0.7	1272	2	Q9F553	Q9F553 escherichia
273	7	0.7	546	2	Q80243	Q80243 mycoplasma	346	7	0.7	1292	3	Q9HDM7	Q9HDM7 schizosacch
274	7	0.7	548	2	Q939K6	Q939K6 propionibac	347	7	0.7	1334	3	Q9Y706	Q9Y706 schizosacch
275	7	0.7	551	4	Q16630	Q16630 homo sapien	348	7	0.7	1349	5	Q9NE31	Q9NE31 leishmania
276	7	0.7	561	16	Q977K3	Q977K3 clostridium	349	7	0.7	1366	5	Q9V8W2	Q9V8W2 drosophila
277	7	0.7	563	4	Q9HCC0	Q9HCC0 h non-biocl	350	7	0.7	1420	5	Q9NEV7	Q9NEV7 drosophila
278	7	0.7	573	10	Q9J211	Q9J211 leishmania	351	7	0.7	1430	5	Q9W0B0	Q9W0B0 drosophila
279	7	0.7	573	10	Q9LWU1	Q9LWU1 arbidopsis	352	7	0.7	1474	5	Q19106	Q19106 caenorhabd
280	7	0.7	588	4	Q9BM18	Q9BM18 homo sapien	353	7	0.7	1597	4	Q9P2D8	Q9P2D8 homo sapien
281	7	0.7	589	5	Q18658	Q18658 trypanosoma	354	7	0.7	1603	2	Q9KKA6	Q9KKA6 rickettsia
282	7	0.7	591	16	Q9RM57	Q9RM57 delinococcus	355	7	0.7	1643	2	Q9F0P7	Q9F0P7 rickettsia
283	7	0.7	595	3	Q94573	Q94573 schizosacch	356	7	0.7	1643	2	Q9F0P6	Q9F0P6 rickettsia
284	7	0.7	596	16	Q98R33	Q98R33 mycoplasma	357	7	0.7	1732	5	Q09451	Q09451 bonnemaisson
285	7	0.7	600	4	Q961B3	Q961B3 homo sapien	358	7	0.7	1745	10	Q9LIM7	Q9LIM7 oryza sativ
286	7	0.7	605	5	Q9GYD0	Q9GYD0 leishmania	359	7	0.7	1884	4	Q9LID7	Q9LID7 homo sapien
287	7	0.7	605	17	Q58850	Q58850 pyrococcus	360	7	0.7	1899	10	Q9XEG1	Q9XEG1 gossypium h
288	7	0.7	607	16	Q92W96	Q92W96 rhizobium m	361	7	0.7	1931	10	Q9SFV6	Q9SFV6 arbidopsis
289	7	0.7	608	17	Q9UZM6	Q9UZM6 pyrococcus	362	7	0.7	1953	5	Q9VKD0	Q9VKD0 drosophila
290	7	0.7	615	3	Q9P861	Q9P861 arxula aden	363	7	0.7	2093	4	Q15010	Q15010 homo sapien
291	7	0.7	622	13	Q919U5	Q919U5 hippoglossu	364	7	0.7	2149	10	Q9ZM93	Q9ZM93 arbidopsis
292	7	0.7	625	3	Q74922	Q74922 schizosacch	365	7	0.7	2391	16	Q99X42	Q99X42 streptomyce
293	7	0.7	644	5	Q9N4W3	Q9N4W3 caenorhabd	366	7	0.7	2435	5	Q9M433	Q9M433 drosophila
294	7	0.7	647	3	Q9P469	Q9P469 neurospora	367	7	0.7	2571	2	Q87704	Q87704 bacillus su
295	7	0.7	656	5	Q963J2	Q963J2 drosophila	368	7	0.7	2634	2	Q95J22	Q95J22 leishmania
296	7	0.7	664	16	Q92C48	Q92C48 listeria in	369	7	0.7	4685	9	Q9R7E8	Q9R7E8 escherichia
297	7	0.7	672	10	Q9NK74	Q9NK74 drosophila	370	7	0.6	9	2	Q9RRD8	Q9RRD8 erythroglus
298	7	0.7	672	10	Q94IM7	Q94IM7 hordeum vul	371	7	0.6	18	6	Q9MY25	Q9MY25 homo sapien
299	7	0.7	673	10	Q94IM6	Q94IM6 hordeum vul	372	7	0.6	36	7	Q9KX25	Q9KX25 bacterioph
300	7	0.7	678	2	Q9AGM8	Q9AGM8 legionella	373	7	0.6	37	9	Q9S506	Q9S506 bacterioph
301	7	0.7	678	10	Q65709	Q65709 arbidopsis	374	7	0.6	41	7	Q97980	Q97980 homo sapien
302	7	0.7	680	10	Q9M617	Q9M617 brassica ju	375	7	0.6	41	9	Q9T156	Q9T156 bacterioph
303	7	0.7	694	17	Q9YG21	Q9YG21 aeropyrum p	376	7	0.6	41	15	Q9EKA4	Q9EKA4 human immun
304	7	0.7	701	17	Q973X0	Q973X0 sulfolobus	377	7	0.6	44	16	Q98AW7	Q98AW7 rhizobium l
305	7	0.7	709	2	Q52772	Q52772 myxococcus	378	7	0.6	44	5	Q9TXE8	Q9TXE8 caenorhabd
306	7	0.7	742	2	Q9F682	Q9F682 streptococ	379	7	0.6	44	10	Q9ARK1	Q9ARK1 peonia mai
307	7	0.7	752	5	Q21026	Q21026 caenorhabd	380	7	0.6	48	5	Q9GR14	Q9GR14 nematostell
308	7	0.7	752	12	Q98140	Q98140 kaposi's sa	381	7	0.6	50	12	Q91G19	Q91G19 chilo lride

382	6	0.6	51	2	Q9F8M3	Q9F8M3 carboxydoth	455	6	0.6	84	15	Q9YUM1	Q9YUM1 human immun
383	6	0.6	51	12	Q9E8B8	Q9E8B8 gallid herp	456	6	0.6	84	15	Q9WLE6	Q9WLE6 human immun
384	6	0.6	51	12	Q9E8B6	Q9E8B6 gallid herp	457	6	0.6	84	15	Q9YUJ9	Q9YUJ9 human immun
385	6	0.6	53	2	P77410	P77410 escherichia	458	6	0.6	84	15	Q9YUJ6	Q9YUJ6 human immun
386	6	0.6	53	5	O19154	O19154 caenorhabdi	459	6	0.6	84	15	Q9PQ01	Q9PQ01 human immun
387	6	0.6	56	11	Q9R082	Q9R082 mus musculu	460	6	0.6	84	16	Q9ZVP8	Q9ZVP8 human immun
388	6	0.6	57	2	O05320	O05320 mus musculu	461	6	0.6	85	11	Q882P8	Q882P8 thizobium m
389	6	0.6	57	12	P89008	P89008 mumps virus	462	6	0.6	85	11	Q882P6	Q882P6 mus musculu
390	6	0.6	57	12	P89010	P89010 mumps virus	463	6	0.6	86	2	Q9REF3	Q9REF3 human immun
391	6	0.6	57	12	P89013	P89013 mumps virus	464	6	0.6	86	10	Q38768	Q38768 bradyrhizob
392	6	0.6	57	12	Q98349	Q98349 mumps virus	465	6	0.6	86	12	Q71071	Q71071 avena sativ
393	6	0.6	57	12	Q98350	Q98350 mumps virus	466	6	0.6	86	15	Q9WQ62	Q9WQ62 human aden
394	6	0.6	57	12	Q9QJX0	Q9QJX0 mumps virus	467	6	0.6	86	15	Q994N0	Q994N0 human immun
395	6	0.6	57	12	Q9QJW9	Q9QJW9 mumps virus	468	6	0.6	87	10	Q94EVL	Q94EVL oryza sativ
396	6	0.6	57	12	Q9QJW8	Q9QJW8 mumps virus	469	6	0.6	87	13	Q90692	Q90692 gallus gall
397	6	0.6	57	12	Q98351	Q98351 mumps virus	470	6	0.6	88	15	Q9WQ64	Q9WQ64 human immun
398	6	0.6	61	2	Q93PR1	Q93PR1 bradyrhizob	471	6	0.6	89	2	Q33M16	Q33M16 human immun
399	6	0.6	61	9	Q9AYZ6	Q9AYZ6 bacterioph	472	6	0.6	89	5	Q9VQC5	Q9VQC5 azotobacter
400	6	0.6	61	16	Q9JUT4	Q9JUT4 neisseria m	473	6	0.6	89	8	Q63701	Q63701 drosophila r
401	6	0.6	64	2	Q9RPL9	Q9RPL9 streptococc	474	6	0.6	89	15	Q98XY4	Q98XY4 human immun
402	6	0.6	64	2	Q9FDN2	Q9FDN2 streptococc	475	6	0.6	89	15	Q9DVP6	Q9DVP6 human immun
403	6	0.6	64	2	Q9F0G5	Q9F0G5 streptococc	476	6	0.6	89	16	Q31715	Q31715 bacillus su
404	6	0.6	64	2	Q9XAL4	Q9XAL4 streptomyce	477	6	0.6	90	16	Q921X1	Q921X1 rickettsia
405	6	0.6	65	16	Q9ZJL5	Q9ZJL5 rickettsia	478	6	0.6	91	10	Q9LE62	Q9LE62 oryza sativ
406	6	0.6	66	10	Q23929	Q23929 flaveria pr	479	6	0.6	91	11	Q9CTE0	Q9CTE0 mus musculu
407	6	0.6	66	10	Q23932	Q23932 flaveria tr	480	6	0.6	91	12	Q9PYR8	Q9PYR8 xestia c-ni
408	6	0.6	66	12	Q9WML7	Q9WML7 human polio	481	6	0.6	91	15	Q9PKX5	Q9PKX5 human immun
409	6	0.6	66	12	Q9WML5	Q9WML5 human polio	482	6	0.6	91	15	Q9PKX4	Q9PKX4 human immun
410	6	0.6	67	5	Q9NM93	Q9NM93 leishmania	483	6	0.6	91	15	Q9QAU9	Q9QAU9 human immun
411	6	0.6	68	16	Q32072	Q32072 bacillus su	484	6	0.6	91	15	Q9QAU5	Q9QAU5 human immun
412	6	0.6	68	16	Q97NV5	Q97NV5 streptococc	485	6	0.6	91	15	Q9Q4U1	Q9Q4U1 human immun
413	6	0.6	70	7	Q95194	Q95194 stizostedio	486	6	0.6	91	15	Q98XX8	Q98XX8 human immun
414	6	0.6	70	7	Q95192	Q95192 stizostedio	487	6	0.6	91	15	Q98XX2	Q98XX2 human immun
415	6	0.6	70	7	Q95191	Q95191 stizostedio	488	6	0.6	91	16	Q9PE14	Q9PE14 xylella fas
416	6	0.6	70	7	Q95188	Q95188 stizostedio	489	6	0.6	91	16	Q9KVR2	Q9KVR2 stizostedio
417	6	0.6	70	7	Q95183	Q95183 stizostedio	490	6	0.6	92	2	Q51557	Q51557 vibrio chol
418	6	0.6	71	2	Q9ZFR3	Q9ZFR3 mastigoclad	491	6	0.6	92	15	Q77705	Q77705 pseudomonas
419	6	0.6	71	4	Q13969	Q13969 homo sapien	492	6	0.6	92	15	Q77709	Q77709 human immun
420	6	0.6	71	16	Q92224	Q92224 rhizobium m	493	6	0.6	92	11	Q77709	Q77709 human immun
421	6	0.6	72	2	Q9AN03	Q9AN03 bradyrhizob	494	6	0.6	93	11	Q89082	Q89082 mus musculu
422	6	0.6	72	5	Q9XZL5	Q9XZL5 comus stria	495	6	0.6	93	15	Q75613	Q75613 human immun
423	6	0.6	72	17	Q97W34	Q97W34 sulfolobus	496	6	0.6	94	2	Q9KRF1	Q9KRF1 clostridium
424	6	0.6	72	17	Q974G6	Q974G6 sulfolobus	497	6	0.6	94	15	Q91X08	Q91X08 human immun
425	6	0.6	73	2	Q9RJ22	Q9RJ22 streptomyce	498	6	0.6	95	17	Q27927	Q27927 methanother
426	6	0.6	76	16	Q9KAK0	Q9KAK0 bacillus ha	499	6	0.6	96	12	P87596	P87596 cowpox viru
427	6	0.6	77	2	Q9ZAG1	Q9ZAG1 shigella so	500	6	0.6	96	15	Q9DVQ9	Q9DVQ9 human immun
428	6	0.6	77	3	Q96W00	Q96W00 verticilliu	501	6	0.6	96	15	Q9DVP3	Q9DVP3 human immun
429	6	0.6	77	3	Q96WF6	Q96WF6 torrubella	502	6	0.6	97	10	Q93WA2	Q93WA2 arabisdopsis
430	6	0.6	77	3	Q96WA4	Q96WA4 escovopsis	503	6	0.6	97	15	Q9QAU7	Q9QAU7 human immun
431	6	0.6	77	3	Q96WA3	Q96WA3 escovopsis	504	6	0.6	97	15	Q9QAU6	Q9QAU6 human immun
432	6	0.6	77	4	Q9NS11	Q9NS11 homo sapien	505	6	0.6	97	15	Q98XY6	Q98XY6 human immun
433	6	0.6	77	16	Q9PGF7	Q9PGF7 xylella fas	506	6	0.6	97	16	Q51726	Q51726 borella bu
434	6	0.6	78	11	Q9QW11	Q9QW11 mus sp. a-r	507	6	0.6	98	10	Q9FHV4	Q9FHV4 arabisdopsis
435	6	0.6	79	2	Q9AEV9	Q9AEV9 lactococcus	508	6	0.6	98	15	Q9YOV8	Q9YOV8 human immun
436	6	0.6	79	4	Q9H566	Q9H566 homo sapien	509	6	0.6	98	15	Q9YOV4	Q9YOV4 human immun
437	6	0.6	79	12	Q68100	Q68100 human cytom	510	6	0.6	98	15	Q91Q06	Q91Q06 human immun
438	6	0.6	79	15	Q98Y00	Q98Y00 human immun	511	6	0.6	98	15	Q9YQ76	Q9YQ76 human immun
439	6	0.6	79	16	Q25449	Q25449 helicobacte	512	6	0.6	98	15	Q9Q4V7	Q9Q4V7 human immun
440	6	0.6	79	16	Q9ZL89	Q9ZL89 helicobacte	513	6	0.6	98	15	Q9Q4U4	Q9Q4U4 human immun
441	6	0.6	80	2	Q9EWN2	Q9EWN2 streptomyce	514	6	0.6	98	15	Q9Q4T4	Q9Q4T4 human immun
442	6	0.6	80	12	Q9YVZ7	Q9YVZ7 melaniopus	515	6	0.6	98	15	Q91X04	Q91X04 human immun
443	6	0.6	80	15	Q9PFX3	Q9PFX3 human immun	516	6	0.6	99	2	Q12935	Q12935 mycobacteri
444	6	0.6	81	3	Q9I185	Q9I185 schizosacch	517	6	0.6	99	8	Q9WJMO	Q9WJMO scenedesmus
445	6	0.6	81	16	Q9CKG4	Q9CKG4 pasteurella	518	6	0.6	99	11	Q9CVU7	Q9CVU7 mus musculu
446	6	0.6	82	8	Q9TH23	Q9TH23 nicotiana r	519	6	0.6	99	15	Q10177	Q10177 human immun
447	6	0.6	82	8	Q9X1X8	Q9X1X8 nicotiana t	520	6	0.6	99	15	Q9PKX7	Q9PKX7 human immun
448	6	0.6	82	10	Q9ZPM0	Q9ZPM0 arabisdopsis	521	6	0.6	99	15	Q9PKX6	Q9PKX6 human immun
449	6	0.6	82	16	Q91625	Q91625 pseudomonas	522	6	0.6	99	15	Q9W8A9	Q9W8A9 human immun
450	6	0.6	83	15	Q9YUJ8	Q9YUJ8 human immun	523	6	0.6	99	15	Q9E9W3	Q9E9W3 human immun
451	6	0.6	83	15	Q9Q4U6	Q9Q4U6 human immun	524	6	0.6	99	15	Q9E9W1	Q9E9W1 human immun
452	6	0.6	84	12	Q91F32	Q91F32 cydia pomon	525	6	0.6	99	15	Q9E9W0	Q9E9W0 human immun
453	6	0.6	84	15	Q9YUM4	Q9YUM4 human immun	526	6	0.6	99	15	Q10764	Q10764 human immun
454	6	0.6	84	15	Q9YUM3	Q9YUM3 human immun	527	6	0.6	99	15	Q12868	Q12868 human immun

528	6	0.6	99	15	010766	010766	human	immun	601	038875	038875	human	immun
529	6	0.6	99	15	010767	010767	human	immun	602	038878	038878	human	immun
530	6	0.6	99	15	012869	012869	human	immun	603	038879	038879	human	immun
531	6	0.6	99	15	010770	010770	human	immun	604	038880	038880	human	immun
532	6	0.6	99	15	010774	010774	human	immun	605	038881	038881	human	immun
533	6	0.6	99	15	010776	010776	human	immun	606	038882	038882	human	immun
534	6	0.6	99	15	010779	010779	human	immun	607	038883	038883	human	immun
535	6	0.6	99	15	010780	010780	human	immun	608	038884	038884	human	immun
536	6	0.6	99	15	012870	012870	human	immun	609	038885	038885	human	immun
537	6	0.6	99	15	036753	036753	human	immun	610	038887	038887	human	immun
538	6	0.6	99	15	036754	036754	human	immun	611	038888	038888	human	immun
539	6	0.6	99	15	038711	038711	human	immun	612	038889	038889	human	immun
540	6	0.6	99	15	038712	038712	human	immun	613	038890	038890	human	immun
541	6	0.6	99	15	038713	038713	human	immun	614	038891	038891	human	immun
542	6	0.6	99	15	038714	038714	human	immun	615	038892	038892	human	immun
543	6	0.6	99	15	038715	038715	human	immun	616	038893	038893	human	immun
544	6	0.6	99	15	038774	038774	human	immun	617	038894	038894	human	immun
545	6	0.6	99	15	038780	038780	human	immun	618	038895	038895	human	immun
546	6	0.6	99	15	038788	038788	human	immun	619	038896	038896	human	immun
547	6	0.6	99	15	038790	038790	human	immun	620	038935	038935	human	immun
548	6	0.6	99	15	038794	038794	human	immun	621	038937	038937	human	immun
549	6	0.6	99	15	038797	038797	human	immun	622	038945	038945	human	immun
550	6	0.6	99	15	038798	038798	human	immun	623	038946	038946	human	immun
551	6	0.6	99	15	038799	038799	human	immun	624	038955	038955	human	immun
552	6	0.6	99	15	038800	038800	human	immun	625	038965	038965	human	immun
553	6	0.6	99	15	038801	038801	human	immun	626	038969	038969	human	immun
554	6	0.6	99	15	038802	038802	human	immun	627	038970	038970	human	immun
555	6	0.6	99	15	038803	038803	human	immun	628	039046	039046	human	immun
556	6	0.6	99	15	038807	038807	human	immun	629	039050	039050	human	immun
557	6	0.6	99	15	038808	038808	human	immun	630	039052	039052	human	immun
558	6	0.6	99	15	038810	038810	human	immun	631	039053	039053	human	immun
559	6	0.6	99	15	038811	038811	human	immun	632	039055	039055	human	immun
560	6	0.6	99	15	038812	038812	human	immun	633	039058	039058	human	immun
561	6	0.6	99	15	038813	038813	human	immun	634	039059	039059	human	immun
562	6	0.6	99	15	038814	038814	human	immun	635	039062	039062	human	immun
563	6	0.6	99	15	038815	038815	human	immun	636	039063	039063	human	immun
564	6	0.6	99	15	038816	038816	human	immun	637	039065	039065	human	immun
565	6	0.6	99	15	038817	038817	human	immun	638	039067	039067	human	immun
566	6	0.6	99	15	038822	038822	human	immun	639	039068	039068	human	immun
567	6	0.6	99	15	038824	038824	human	immun	640	039070	039070	human	immun
568	6	0.6	99	15	038828	038828	human	immun	641	039071	039071	human	immun
569	6	0.6	99	15	038829	038829	human	immun	642	039072	039072	human	immun
570	6	0.6	99	15	038830	038830	human	immun	643	039074	039074	human	immun
571	6	0.6	99	15	038832	038832	human	immun	644	039075	039075	human	immun
572	6	0.6	99	15	038833	038833	human	immun	645	039076	039076	human	immun
573	6	0.6	99	15	038836	038836	human	immun	646	039079	039079	human	immun
574	6	0.6	99	15	038838	038838	human	immun	647	039080	039080	human	immun
575	6	0.6	99	15	038840	038840	human	immun	648	039081	039081	human	immun
576	6	0.6	99	15	038841	038841	human	immun	649	039083	039083	human	immun
577	6	0.6	99	15	038842	038842	human	immun	650	039084	039084	human	immun
578	6	0.6	99	15	038844	038844	human	immun	651	039086	039086	human	immun
579	6	0.6	99	15	038845	038845	human	immun	652	039091	039091	human	immun
580	6	0.6	99	15	038847	038847	human	immun	653	039092	039092	human	immun
581	6	0.6	99	15	038848	038848	human	immun	654	039093	039093	human	immun
582	6	0.6	99	15	038850	038850	human	immun	655	039094	039094	human	immun
583	6	0.6	99	15	038853	038853	human	immun	656	039096	039096	human	immun
584	6	0.6	99	15	038854	038854	human	immun	657	039097	039097	human	immun
585	6	0.6	99	15	038855	038855	human	immun	658	039100	039100	human	immun
586	6	0.6	99	15	038856	038856	human	immun	659	039101	039101	human	immun
587	6	0.6	99	15	038857	038857	human	immun	660	039102	039102	human	immun
588	6	0.6	99	15	038858	038858	human	immun	661	039103	039103	human	immun
589	6	0.6	99	15	038859	038859	human	immun	662	039104	039104	human	immun
590	6	0.6	99	15	038860	038860	human	immun	663	039105	039105	human	immun
591	6	0.6	99	15	038861	038861	human	immun	664	039138	039138	human	immun
592	6	0.6	99	15	038862	038862	human	immun	665	039142	039142	human	immun
593	6	0.6	99	15	038863	038863	human	immun	666	039144	039144	human	immun
594	6	0.6	99	15	038864	038864	human	immun	667	039152	039152	human	immun
595	6	0.6	99	15	038865	038865	human	immun	668	039156	039156	human	immun
596	6	0.6	99	15	038868	038868	human	immun	669	039159	039159	human	immun
597	6	0.6	99	15	038869	038869	human	immun	670	039160	039160	human	immun
598	6	0.6	99	15	038872	038872	human	immun	671	092144	092144	human	immun
599	6	0.6	99	15	038873	038873	human	immun	672	090650	090650	human	immun
600	6	0.6	99	15	038874	038874	human	immun	673	090651	090651	human	immun


```

966 6 0.6 99 15 Q9DLX4 Q9DLX4 human human
967 6 0.6 99 15 Q9DLX3 Q9DLX3 human human
968 6 0.6 99 15 Q9DLX2 Q9DLX2 human human
969 6 0.6 99 15 Q9DLX1 Q9DLX1 human human
970 6 0.6 99 15 Q9DLX0 Q9DLX0 human human
971 6 0.6 99 15 Q9DLX9 Q9DLX9 human human
972 6 0.6 99 15 Q9DLX8 Q9DLX8 human human
973 6 0.6 99 15 Q9DLX7 Q9DLX7 human human
974 6 0.6 99 15 Q9DLX6 Q9DLX6 human human
975 6 0.6 99 15 Q9DLX5 Q9DLX5 human human
976 6 0.6 99 15 Q9DLX4 Q9DLX4 human human
977 6 0.6 99 15 Q9DLX3 Q9DLX3 human human
978 6 0.6 99 15 Q9DLX2 Q9DLX2 human human
979 6 0.6 99 15 Q9DLX1 Q9DLX1 human human
980 6 0.6 99 15 Q9DLX0 Q9DLX0 human human
981 6 0.6 99 15 Q9DLW9 Q9DLW9 human human
982 6 0.6 99 15 Q9DLW8 Q9DLW8 human human
983 6 0.6 99 15 Q9DLW7 Q9DLW7 human human
984 6 0.6 99 15 Q9DLW6 Q9DLW6 human human
985 6 0.6 99 15 Q9DLW5 Q9DLW5 human human
986 6 0.6 99 15 Q9DLW4 Q9DLW4 human human
987 6 0.6 99 15 Q9DLW3 Q9DLW3 human human
988 6 0.6 99 15 Q9DLW2 Q9DLW2 human human
989 6 0.6 99 15 Q9DLW1 Q9DLW1 human human
990 6 0.6 99 15 Q9DLW0 Q9DLW0 human human
991 6 0.6 99 15 Q9DLV9 Q9DLV9 human human
992 6 0.6 99 15 Q9DLV8 Q9DLV8 human human
993 6 0.6 99 15 Q9DLV7 Q9DLV7 human human
994 6 0.6 99 15 Q9DLV6 Q9DLV6 human human
995 6 0.6 99 15 Q9DLV5 Q9DLV5 human human
996 6 0.6 99 15 Q9DLV4 Q9DLV4 human human
997 6 0.6 99 15 Q9DLV3 Q9DLV3 human human
998 6 0.6 99 15 Q9DLV2 Q9DLV2 human human
999 6 0.6 99 15 Q9DLV1 Q9DLV1 human human
1000 6 0.6 99 15 Q9DLU9 Q9DLU9 human human

```

ALIGNMENTS

```

RESULT 1
ID Q9HCD4 PRELIMINARY: PRT: 905 AA.
AC Q9HCD4:
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)
DE KIA1638 PROTEIN (FRAGMENT).
CN KIA1638.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-20450683; PubMed-10997877;
RA Nagase T., Kikuno R., Nakayama M., Hirosewa M., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes.
RT XVII. The complete sequences of 100 new cDNA clones from brain which
RT code for large proteins in vitro.";
RL DNA Res. 7:273-281(2000).
DR EMBL; AB046858; BAB13464.1;
DR InterPro; IPR000547; Clathrin_repeat.
DR SMART; SM00299; CLH; 1.
FT NON_TER 1
SQ SEQUENCE 905 AA: 102450 MW: PD370996F00FE1BF CRC64;

```

```

Query Match 46.4%; Score 498; DB 4; Length 905;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 698; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

OY 375 ALSTORGSLHVFELTKLPITIGDACSRIAYLTSLLEVTVANPVEGELPTIVSDVEPNFVA 434

```

Db 1 ALSTORGSLHVFELTKLPITIGDACSRIAYLTSLLEVTVANPVEGELPTIVSDVEPNFVA 60
435 VGLYHLAVGMNRAWFYVGENAVKKLDMETLGVVASICLHSDPAALFEKVOHLIE 494
61 VGLYHLAVGMNRAWFYVGENAVKKLDMETLGVVASICLHSDPAALFEKVOHLIE 120
495 SEILDAGEERETRLPPAVDDKCRILCHALTSDFLLYGDTGVVOYVEIEDMOFVNDYBNP 554
121 SEILDAGEERETRLPPAVDDKCRILCHALTSDFLLYGDTGVVOYVEIEDMOFVNDYBNP 180
555 VSVKRIFFDPNKTRLFVFEDEKSGFVYCPVNDATYEIYDFSTFYGVMENPMKGVFI 614
181 VSVKRIFFDPNKTRLFVFEDEKSGFVYCPVNDATYEIYDFSTFYGVMENPMKGVFI 240
615 AYDDDKVTVYVHKDTIGAKVILAGSTKVPFAHPLLYNNELTCOTQSGKVNNTYST 674
241 AYDDDKVTVYVHKDTIGAKVILAGSTKVPFAHPLLYNNELTCOTQSGKVNNTYST 300
675 HGFSLNLKDXGPDDELRLPMLAHNMLKRFSDAWEMCRILNDEAANMLARACLHMEVEFA 734
301 HGFSLNLKDXGPDDELRLPMLAHNMLKRFSDAWEMCRILNDEAANMLARACLHMEVEFA 360
735 IYVYRIRIGVGVVMSLEQIKGIEDYNLLAGHLAMFTNDYNLAQDLYLASSCPIALLEMR 794
361 IYVYRIRIGVGVVMSLEQIKGIEDYNLLAGHLAMFTNDYNLAQDLYLASSCPIALLEMR 420
795 DLQHMDSALQLAKHLAPQIPIISKEVAIQLEFAGDYVNALAHYKGTGDKREDECL 854
421 DLQHMDSALQLAKHLAPQIPIISKEVAIQLEFAGDYVNALAHYKGTGDKREDECL 480
855 AGVAQMSIRMGDIRRGVNAQALKHSRVLRKDCGAILLENNKQFSEAAQLYKGLYDKAAS 914
481 AGVAQMSIRMGDIRRGVNAQALKHSRVLRKDCGAILLENNKQFSEAAQLYKGLYDKAAS 540
915 VYIRSKNNAKVGDLPLHYSSPKIHLOYAKARADRYEAVAYANAKOMOSVIRIYLDH 974
541 VYIRSKNNAKVGDLPLHYSSPKIHLOYAKARADRYEAVAYANAKOMOSVIRIYLDH 600
975 LNNPEKAVNIARETQSLDGAKVAVRFLQDYGSAIOFLVMSKCNNEAFLAQOHKME 1034
601 LNNPEKAVNIARETQSLDGAKVAVRFLQDYGSAIOFLVMSKCNNEAFLAQOHKME 660
OY 1035 IYADIIIGSEDTTNEGYSTALYFEGEKRYLQAGKFFLLCG 1074
661 IYADIIIGSEDTTNEGYSTALYFEGEKRYLQAGKFFLLCG 700

```

```

RESULT 2
ID Q9H5S0 PRELIMINARY: PRT: 481 AA.
AC Q9H5S0:
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)
DE CDNA: FLJ23127 F1S, CLONE LMG08350.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE-LONG;
RA Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T., Matsumura K.,
RA Kawakami T., Mizuno T., Morinaga M., Tanigami A., Fujiwara T., Ono T.,
RA Yamada K., Fujii Y., Ozaki K., Hirose M., Ohmori Y., Ota T., Suzuki Y.,
RA Odayashi M., Nishi T., Shibahara T., Tanaka T., Nakamura Y.,
RA Isogai T., Sugano S.;
RT "NEBO human cDNA sequencing project.";
RL Submitted (AUG-2000) to the EMBL/Genbank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 3 WD REPEATS (TRP-ASP DOMAINS).
DR EMBL; AK026780; BAB15550.1;
DR InterPro; IPR001680; WD40.

```

DR Pfam: PF00400; WD40; 3.
 DR SMART: SM00320; WD40; 5.
 KM Repeat: WD repeat.
 SQ SEQUENCE 481 AA; 53542 MW; B1C7AE099C581F02 CRC64;

Query Match 25.1%; Score 270; DB 4; Length 481;

Best Local Similarity 100.0%; Pred. No. 36-288;
 Matches 270; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 223 NMDFLLKMDRTSAASMSIVLGKTLFFLNINPDNPADLFQDFGNIYCYMYGD 282
 DB 183 NMDFLLKMDRTSAASMSIVLGKTLFFLNINPDNPADLFQDFGNIYCYMYGD 242
 QY 283 GRIGFSCGFYVYSTRHGEIQRNRHNDLTSIAVSQFLNVAACGDCICIKD 342
 DB 243 GRIGFSCGFYVYSTRHGEIQRNRHNDLTSIAVSQFLNVAACGDCICIKD 302
 QY 343 LVLDKMYVILNDEENKGLGTLSDGQLALSTORGLHVEFLTKPLIGDASTRIA 402
 DB 303 LVLDKMYVILNDEENKGLGTLSDGQLALSTORGLHVEFLTKPLIGDASTRIA 362
 QY 403 YLTSLEEVYVAVPVEGLPTTVSDVDPNRYAVGLYHLAVGMNRRANFYVGENAVKLLK 462
 DB 363 YLTSLEEVYVAVPVEGLPTTVSDVDPNRYAVGLYHLAVGMNRRANFYVGENAVKLLK 422
 QY 463 DMEYLGTVASICLHSDYAALFEGKVQLHL 492
 DB 423 DMEYLGTVASICLHSDYAALFEGKVQLHL 452

RESULT 3

046018 PRELIMINARY; PRT; 519 AA.
 AC 046018;

DT 01-JUN-1998 (TREMBLrel. 06, Created)
 DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE ZK520.1 PROTEIN.
 CN ZK520.1.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Pelodierinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Steward C.A.;
 RN Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99069613; PubMed=9851916;
 RA none;
 RT "Genome sequence of the nematode C. elegans: A platform for
 RT investigating biology."
 DR EMBL: Z92822; CAB07299.1; Science 282:2012-2018(1998).
 SQ SEQUENCE 519 AA; 59152 MW; 38F80CA32D06E02D CRC64;

Query Match 0.8%; Score 9; DB 5; Length 519;
 Best Local Similarity 100.0%; Pred. No. 2.2;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1009 SAIOFLVWS 1017
 DB 293 SAIOFLVWS 301

RESULT 4
 Q9GY76 PRELIMINARY; PRT; 706 AA.
 AC Q9GY76;
 DT 01-MAR-2001 (TREMBLrel. 16, Created)

DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE PROBABLE HYPOTHETICAL 11.6 KDA PROTEIN.
 GN LM12.54.
 OS Leishmania major.
 OC Eukaryota; Euzoenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
 OX NCBI_TaxID=5664;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FRIDLIN;
 RA Murphy L.; Quail M.; Harris D.; Rajandream M.; Ivens A.; Barrell B.;
 RA Oliver K.;
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AL390114; CAC01977.2;
 SQ SEQUENCE 706 AA; 72716 MW; 29199F2AA4CC64FD CRC64;

Query Match 0.8%; Score 9; DB 5; Length 706;
 Best Local Similarity 100.0%; Pred. No. 2.9;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 368 TDDGQLLAL 376
 DB 388 TDDGQLLAL 396

RESULT 5

0995R1 PRELIMINARY; PRT; 120 AA.
 ID 0995R1;

DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE GAG POLYPEPTIDE (FRAGMENT).
 GN GAG.
 OS Human immunodeficiency virus type 1.
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11676;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=M6;
 RA Virell D.L.;
 RT "Immunological and virological characteristics of HIV infected
 RT individuals in different risk groups in Madras, India."
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF250278; AK3492.1;
 DR InterPro: IPR00007L; Retroviral_gag_p17.
 DR Pfam: PF00540; gag_p17; 1.
 DR PRINTS: PR00234; HIVMATRIX.
 KW AIDS; Core protein; Polypeptide.
 FT NON_TER 1
 FT NON_TER 1
 FT SEQUENCE 120 AA; 13544 MW; 6D4B7597A6ACFB3C CRC64;

Query Match 0.7%; Score 8; DB 15; Length 120;
 Best Local Similarity 100.0%; Pred. No. 7;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 942 AKAKEADG 949
 DB 90 AKAKEADG 97

RESULT 6
 Q84636 PRELIMINARY; PRT; 176 AA.
 AC 084636;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE A322L PROTEIN.
 GN A322L.

```

OS Parametium bursaria chlorella virus 1 (PBCV-1).
OC Viruses: dsDNA viruses, no RNA stage; Phycodnaviridae; Phycodnavirus.
OX NCB1_TaxID=10506;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96187795; PubMed=8614977;
RA Lu Z., Li Y., Que Q., Kutish G.F., Rock D.L., Van Etten J.L.;
RT "Analysis of 94 kb of the chlorella virus PBCV-1 330-Kb genome: map
RL positions 88 to 182."
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=20013326; PubMed=10544099;
RA Kaiser A., Vollmert M., Tholl D., Graves M.V., Gurnon J.R., Xing W.,
RA Lisee A.D., Mickerson K.W., Van Etten J.L.;
RT "Chlorella virus PBCV-1 encodes a functional homospesmidine
RN synthase."
RL Virology 263:254-262(1999).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=20478054; PubMed=11021991;
RA Sun L., Gurnon J.R., Adams B.J., Graves M.V., Van Etten J.L.;
RT "Characterization of a beta-1,3-glucanase encoded by chlorella virus
RN PBCV-1."
RL Virology 276:27-36(2000).
RN [4]
RP SEQUENCE FROM N.A.
RA Van Etten J.L.;
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RA Van Etten J.L.;
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA Van Etten J.L.;
RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
RN [7]
RP SEQUENCE FROM N.A.
RA Graves M.V., Van Etten J.L.;
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
RN [8]
RP SEQUENCE FROM N.A.
RA Graves M.V., Van Etten J.L.;
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
RN [9]
RP SEQUENCE FROM N.A.
RA Gurnon J.R., Graves M.V., Van Etten J.L.;
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: U42580; AAC36690.1;
DR InterPro: IPR003015; HLH_MYC.
DR PROSITE: PS00038; HELIX_LOOP_HELIX; UNKNOWN_1.
SQ SEQUENCE 176 AA; 20039 MW; 8CA9FBC4C16A2F96 CRC64;

Query Match 0.7%; Score 8; DB 12; Length 176;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 628 KDTIOGAK 635
DB 103 KDTIOGAK 110

RESULT 7
ID 09HQ87 PRELIMINARY; PRT; 241 AA.
AC 09HQ87;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DE 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
GN VNG1276C.

```

```

OS Halobacterium sp. (strain NRC-1).
OC Archaea; Euryarchaeota; Halobacteriales; Halobacteriaceae;
OC Halobacterium.
OX NCB1_TaxID=64091;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20504483; PubMed=11016950;
RA Ng W.V., Kennedy S.P., Mahairas G.G., Bergquist B., Pan M.,
RA Shukla H.D., Laskey S.R., Baliga N.S., Thorsson V., Shroga J.,
RA Swartzell S., Weir D., Hall J., Dahl T.A., Welt R., Goo Y.A.,
RA Leithausen B., Keller K., Cruz R., Danson M.J., Hough D.W.,
RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,
RA Isenbarger T.A., Beck R.F., Pohlshöder M., Spudis J.L., Jung K.-H.,
RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA Eberhardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.;
RT "Genome sequence of Halobacterium species NRC-1."
RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=2050551; PubMed=11016950;
RA InterPro: IPR002140; UPEF0023.
DR Pfam: PF01172; UPEF0023; 1.
DR ProDom: PD009796; UPEF0023; 1.
KW Complete proteome.
SQ SEQUENCE 241 AA; 26745 MW; 44513B24D4C79D31 CRC64;

```

```

Query Match 0.7%; Score 8; DB 17; Length 241;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 788 AALEMRD 795
DB 27 AALEMRD 34

RESULT 8
ID 098FH3 PRELIMINARY; PRT; 279 AA.
AC 098FH3;
DT 01-OCT-2001 (TREMBlrel. 18, Created)
DT 01-OCT-2001 (TREMBlrel. 18, Last sequence update)
DT 01-OCT-2001 (TREMBlrel. 18, Last annotation update)
DE STREPTOTHRICIN-ACETYL-TRANSFERASE.
GN ML3776.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCB1_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Idegawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RN Mesorhizobium loti."
RL DNA Res. 7:331-338(2000).
DR EMBL: AP003002; BAB50594.1;
DR InterPro: IPR000073; Abhydrolase.
DR InterPro: IPR003089; AB_hydrolase.
DR InterPro: IPR000379; Est_lip_chloest_actsite.
DR Pfam: PF00561; abhydrolase; 1.
DR PRINTS: PR00111; ABHYDROLASE.
KW Transferase; Complete proteome.
SQ SEQUENCE 279 AA; 30294 MW; 8209435C294073A4 CRC64;

```

```

Query Match 0.7%; Score 8; DB 16; Length 279;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 874 ALKHSRV 881
 |||||
 DB 109 ALKHSRV 116

RESULT 9
 ID 093F73 PRELIMINARY; PRT; 300 AA.
 AC 093F73;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, last annotation update)
 DE MNXG (FRAGMENT).
 GN MNXG.
 OS Bacillus sp. MB-1.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 OC Bacillus/Staphylococcus group; Bacillus.
 OX NCBI_TaxID=161530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MB-1;
 RA Francis C.A., Tebo B.M.;
 RT "Enzymatic manganese(II) oxidation by metabolically-dormant spore of
 RT diverse Bacillus species.";
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF326386; AAL30438.1; -.
 FT NON_TER 1 1
 FT NON_TER 300 300
 SQ SEQUENCE 300 AA; 33592 MW; A058E5CA3801F872 CRC64;

Query Match 0.7%; Score 8; DB 2; Length 300;
 Best Local Similarity 100.0%; Pred. No. 17;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 494 ESEILDQ 501
 |||||
 DB 18 ESEILDQ 25

RESULT 10
 ID 093F70 PRELIMINARY; PRT; 300 AA.
 AC 093F70;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, last annotation update)
 DE MNXG (FRAGMENT).
 GN MNXG.
 OS Bacillus sp. MB-7.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 OC Bacillus/Staphylococcus group; Bacillus.
 OX NCBI_TaxID=161535;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MB-7;
 RA Francis C.A., Tebo B.M.;
 RT "Enzymatic manganese(II) oxidation by metabolically-dormant spore of
 RT diverse Bacillus species.";
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF326389; AAL30441.1; -.
 FT NON_TER 1 1
 FT NON_TER 300 300
 SQ SEQUENCE 300 AA; 33636 MW; DAE6191540BA2761 CRC64;

Query Match 0.7%; Score 8; DB 2; Length 300;
 Best Local Similarity 100.0%; Pred. No. 17;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 494 ESEILDQ 501
 |||||
 DB 18 ESEILDQ 25

RESULT 11
 ID 073906 PRELIMINARY; PRT; 474 AA.
 AC 073906;
 DT 01-AUG-1998 (TREMBlrel. 07, Created)
 DT 01-AUG-1998 (TREMBlrel. 07, last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, last annotation update)
 DE CDS PROTEIN PRECURSOR.
 GN CDS.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=H.B2; TISSUE=THYMUS;
 RX MEDLINE=98250124; PubMed=9590242;
 RT "The structure of avian CDS implies a conserved function.";
 RL J. Immunol. 160:4943-4950(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=B12;
 RX MEDLINE=21332904; PubMed=11439160;
 RA Koskinen R., Salomonsen J., Goodchild M., Bumsstead N., Boyd Y.,
 RA Valinjo O.;
 RT "Structure and chromosomal localization of chicken CDS.";
 RL Scand. J. Immunol. 54:141-145(2001).
 DR EMBL: Y12011; CAAT7239.1; -.
 DR InterPro: IPR001190; SRCR.
 DR Pfam: PF00530; SRCR; 3.
 DR SMART: SM00202; SR; 3.
 DR PROSITE: PSS0287; SRCR_2; 3.
 KW SIGNAL.
 FT SIGNAL 1 22 POTENTIAL.
 FT SIGNAL 474 AA; 50998 MW; 3AF2BA41591AE0D6 CRC64;
 SQ SEQUENCE

Query Match 0.7%; Score 8; DB 13; Length 474;
 Best Local Similarity 100.0%; Pred. No. 25;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 468 GTVASICL 475
 |||||
 DB 358 GTVASICL 365

RESULT 12
 ID 09EV50 PRELIMINARY; PRT; 493 AA.
 AC 09EV50;
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, last annotation update)
 DE CATALASE (EC 1.11.1.6).
 GN KATA.
 OS Staphylococcus xylosus.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 OC Bacillus/Staphylococcus group; Staphylococcus.
 OX NCBI_TaxID=1288;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CURED STRAIN C2A;
 RA Barriere C., Bruckner R., Talon R.;
 RT "Characterization of catalases in Staphylococcus xylosus.";
 RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AJ295151; CAC14836.1; -.
 DR HSSP: P42321; ZCAE.
 DR InterPro: IPR002226; Catalase.

DR Pfam: PF00199; catalase; 1.
 DR PRINTS: PR00067; CATALASE.
 DR Prodom: PD000510; Catalase; 1.
 DR PROSITE: PS00437; CATALASE_1; 1.
 KW Oxidoreductase; Peroxidase.
 SQ SEQUENCE 493 AA; 56874 MW; 900B959518A9404B CRC64;

Query Match 0.7%; Score 8; DB 2; Length 493;
 Best Local Similarity 100.0%; Pred. No. 26;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 310 QARNHKDN 317
 |||||
 DB 266 QARNHKDN 273

RESULT 13
 O9KWI9 PRELIMINARY; PRT; 505 AA.

AC O9KWI9; 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, last sequence update)
 DE 01-DEC-2001 (TREMBLrel. 19, last annotation update)
 DE CATALASE.
 OS Staphylococcus warneri.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 OC Bacillus/Staphylococcus group; Staphylococcus.
 OX NCBI_TaxID=1292;
 RP SEQUENCE FROM N.A.
 RC STRAIN-ISK-1;
 RA Mizuno K., Fukuda D., Kakiyama M., Kohno M., Ha T.L., Sonomoto K.;
 RT "Purification and Gene Cloning of Catalase from Staphylococcus warneri
 ISK-1."
 RL Submitted (JUN-2000) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AB045340; BAA97560.1; -.
 DR HSPF: P42321; ZCAE.
 DR InterPro: IPR002226; Catalase.
 DR Pfam: PF00199; catalase; 1.
 DR PRINTS: PR00067; CATALASE.
 DR Prodom: PD000510; Catalase; 1.
 DR PROSITE: PS00437; CATALASE_1; UNKNOWN_1.
 DR PROSITE: PS00438; CATALASE_2; 1.
 SQ SEQUENCE 505 AA; 58042 MW; 106AF448457EC1F6 CRC64;

Query Match 0.7%; Score 8; DB 2; Length 505;
 Best Local Similarity 100.0%; Pred. No. 27;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 310 QARNHKDN 317
 |||||
 DB 269 QARNHKDN 276

RESULT 14
 O9CD12 PRELIMINARY; PRT; 519 AA.

AC O9CD12; 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, last sequence update)
 DT 01-OCT-2001 (TREMBLrel. 18, last annotation update)
 DE PUTATIVE SECRETED PROTEIN.
 GN MCEIC OR ML2591.
 OS Mycobacterium leprae.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Corynebacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1769;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-TN;
 RX MEDLINE=21128732; PubMed=11234002;

RA Cole S.T., Eiglmeier K., Parkhill J., James K.D., Thomson N.R.,
 RA Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,
 RA Whittall K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R.M., Devlin K., Duthoy S., Feltham T., Fraser A., Hamlin N.,
 RA Holtroyd S., Hornsby T., Jagels K., Lacroix C., Maclean J., Moule S.,
 RA Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,
 RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
 RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
 RA Barrell B.G.;
 RT "Massive gene decay in the leprosy bacillus."
 RL Nature 409:1007-1011(2001).
 DR EMBL: AL583926; CAC32123.1; -.
 DR Leproma: ML2591; -.
 DR InterPro: IPR003399; Mce.
 DR InterPro: IPR002965; P_rich_extensn.
 DR Pfam: PF02470; mce; 1.
 DR PRINTS: PR01217; PRICHEXTENS.
 KW Complete proteome.
 SQ SEQUENCE 519 AA; 54450 MW; 8A3A6AEFAD619EBD CRC64;

Query Match 0.7%; Score 8; DB 16; Length 519;
 Best Local Similarity 100.0%; Pred. No. 28;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 561 FPDPMGTR 568
 |||||
 DB 346 FPDPMGTR 353

RESULT 15
 O993G4 PRELIMINARY; PRT; 699 AA.

AC O993G4; 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, last annotation update)
 DE ORF45.
 OS California herpesvirus 3.
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Gammaherpesvirinae; Lymphocryptovirus.
 OX NCBI_TaxID=106331;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CJ0149;
 RX MEDLINE=21107697; PubMed=11158621;
 RA Cho Y., Ramer J., Rivallier P., Quink C., Garber R.L., Beier D.R.,
 RA Wang F.;
 RT "An Epstein-Barr-related herpesvirus from marmoset lymphomas."
 RL Proc. Natl. Acad. Sci. U.S.A. 98:1224-1229(2001).
 DR EMBL: AF319782; AAK38254.1; -.
 SQ SEQUENCE 699 AA; 75289 MW; 89B75E925AE64AB1 CRC64;

Query Match 0.7%; Score 8; DB 12; Length 699;
 Best Local Similarity 100.0%; Pred. No. 36;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 409 EYTVANPV 416
 |||||
 DB 398 EYTVANPV 405

Search completed: May 21, 2002, 14:56:10
 Job time: 169 sec

THIS PAGE BLANK (USPTO)